ozlemények

11/1973

TA Számítástechnikai és Automatizálási Kutató Intézet Budapest

MAGYAR TUDOMÁNYOS AKADÉMIA
SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZET

közlemények

1973. december

Felelős kiadó:

Dr. VÁMOS TIBOR
igazgató

Technikai szerkesztő:
RÉVÉSZ GYÖRGYI

MTA Számítástechnikai és Automatizálási Kutató Intézete

MTA KESZ Sokszorosító. F.v.: Szabó Gyula

TARTALOMJEGYZÉK

MTA Számítástechnikai és Automatizálási Kutató Intézete, **Közlemények** 11. (1973)

EGY GEOMETRIAI SZÉLSŐÉRTÉKFELADAT

Bán Ilona

Elektromágnesek vasmagjának tervezésénél merül fel a következő geometriai szélsőértékfeladat:

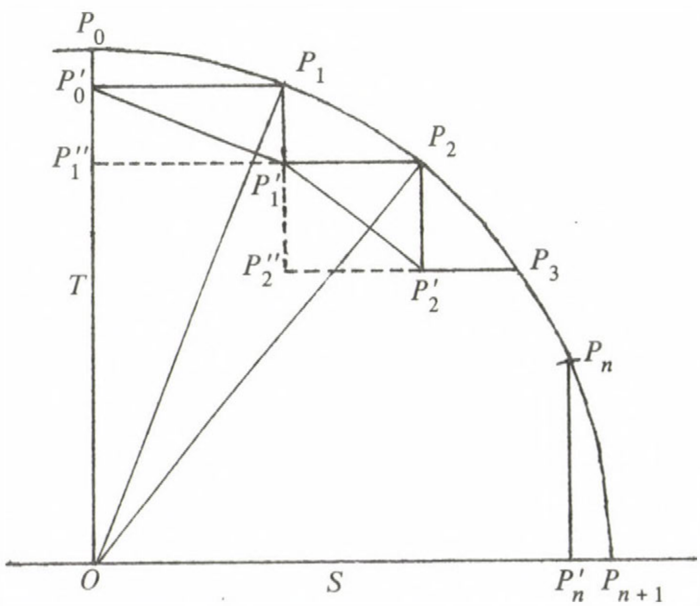
Egy O középpontú K kör kitöltendő előre megadott l számú, párhuzamos elhelyezkedésű, egymással oldalaik mentén érintkező, egymásba nem nyúló, legalább 2 csúcsukkal a köríven lévő téglalapokkal úgy, hogy a téglalapok egyesítéseként adódó L sokszög területe maximális legyen.

Nyilvánvaló, hogy L szimmetrikus az érintkező oldalak S felező merőlegesére. Azt az esetet vizsgáljuk, amikor L szimmetrikus az S-re merőleges, O-n áthaladó T egyenesre is. Igy L-nek $4n$ csúcsa fekszik a körön. Jelölje L_n az ilyen tulajdonságú sokszögek valamelyikét, a maximális területű L_n-t pedig (ha létezik), L_n^*.

L_n^*-nak legalább $4n$ különböző csúcsa van a köríven, hiszen ha L_{n-k}^* csúcsaihoz tőlük különböző csúcsokat veszünk hozzá, akkor a sokszög területe nő, így a maximális terület is. Tehát $L_{n-k}^* < L_n^*$.

S-re és T-re való szimmetria miatt elegendő L_n-nek azt az N_n részét tekinteni, amely az S és T által határolt egyik N negyedkörbe esik. N_n területe változóinak folytonos függvénye, és e változók zárt intervallumon mozognak, ezért N_n területének van legalább egy N_n^* maximuma.

Vezessük be a következő jelöléseket: N_n-nek a negyedköríven levő pontjait T-nek a körívvel való P_0 metszéspontja felöl indulva $P_1, \ldots, P_n$-nek nevezzük. Húzzunk a $P_i (i = 0, \ldots, n)$ pontból T-vel és S-sel párhuzamosokat:



jelöljük T_i-vel illetve S_i-vel. T_i és S_{i+1} metszéspontja P_i', T_{i-1} és S_{i+1} metszéspontja P_i'' legyen. P_{n-1}-nek illetve P_n-nek S-re való vetületét P_n'' illetve P_n'-vel jelöljük.

Vegyük a pontoknak egy olyan elhelyezését, melynél a terület maximális. Ekkor a $P_{i-1}' P_i'' P_i' P_i$ $(i = 1, \ldots, n)$ téglalapok területének is maximálisnak kell lenni. Tegyük fel, hogy ez nem igaz. $P_{i-1}' P_i'' P_i' P_i$ területe nem maximális az $P_{i-1}' P_i''$ és $P_i'' P_i'$ egyenesek által meghatározott tartományban. Keressük meg itt a maximális területű téglalapot és ezt vegyük hozzá az eredeti sokszöghöz. Igy egy nagyobb területű idomot kapunk, ami ellentmond a maximalitásnak.

Tekintsünk egy $P_{i-1}' P_i'' P_i' P_i$ téglalapot. A $P_{i-1}' P_i''$, $P_i'' P_i'$ egyenesekre támaszkodó, vele egyenlő területű téglalapoknak az egyeneseken nem lévő csúcsai egy $P_{i-1}' P_i''$, $P_i'' P_i'$ aszimptotájú hiperbolát írnak le. Ez a hiperbola nem metszheti a $P_{i-1} P_{i+1}$ körívet, hiszen P_i és a tőle különböző metszéspont közötti pontokhoz tartozó téglalap területe nagyobb lenne $P_{i-1}' P_i'' P_i' P_i$-nél, holott a feltevés szerint ez maximális. A hiperbola tehát csak érinti a kört (P_i-ben) és így P_i-ben az érintők közösek. A hiperbola érintője viszont párhuzamos a $P_{i-1}' P_i'' P_i' P_i$ téglalap $P_{i-1}' P_i'$ átlójával. $(i = 1, \ldots, n)$

Ebből következik, hogy P_1 vagy P_n ismeretében a többi csúcspontot már meg lehet szerkeszteni. Legyen pl. megadva a P_1 pont. P_0' -ből a kör P_1-beli érintőjével párhuzamost húzunk. Messük él ezt a P_1-ből az S-re bocsátott merőleges egyenessel. A metszéspontot átfektetett, S-sel párhuzamos egyenes a körívből kimetszi a P_2 pontot. A P_i $(i = 3, 4, \ldots, n)$ pontot úgy kapjuk, hogy a P_{i-1} ponthoz tartozó érintővel a P_{i-2}' pontból párhuzamost húzunk és elmetszük ezt a P_{i-1}-n áthaladó S-re merőleges egyenessel. A metszésponton átfektetett, S-sel párhuzamos egyenes a körívből kimetszi a P_i pontot.

A szerkesztés monotonitásából adódik, hogy egyetlen maximális területű N_n^* sokszög létezik. Megmutatjuk, hogy N_n^* szimmetrikus a negyedkör X szimmetriatengelyére.

Tegyük fel, hogy nem szimmetrikus. Vegyük N_n^*-nak X-re való tükrözését. Nyilván, ennek a területe is maximális, ami ellentmond annak, hogy egyetlen maximum létezik.

Beláttuk tehát, hogy létezik egyetlen maximális területű, n csúcspontú N_n^*, ez szimmetrikus X-re és P_1 vagy P_n ismeretében meg tudjuk szerkeszteni. N_n^*-t S-re és T-re tükrözve megkapjuk L_n^*-t.

A fentiek alapján többféle algoritmus is készíthető a probléma megoldására. Pl. a szimmetria miatt páratlan n esetén az $\frac{n+1}{2}$ -edik pontnak X-re kell esnie, páros n esetén pedig $\frac{n}{2} - 1$ és $\frac{n}{2}$ pontok X-től egyenlő távolságra vannak. Ha páratlan n esetén az $\frac{n+1}{2}$ -edik pont az optimálistól negatív irányba esik, P_1-et pozitív irányba kell mozgatni és fordítva. Páros n esetén is hasonló meggondolás alapján igazíthatjuk be a pontokat és előre megadott pontosságig számolhatjuk a pontok helyét. Ennél egyszerűbb az az algoritmus, ami csak azt használja ki, hogy P_n-ből az előző pontok képzésével megegyező módon kiszámított P_{n+1} pontnak S-

re kell esni n akár páros, akár páratlan.

MEGJEGYZÉSEK

1) Ha nem tesszük fel a T-re való szimmetriát, akkor is tudunk algoritmust adni L_n^* félköríven elhelyezkedő csúcsainak a meghatározására. Ha a félkörben fekvő sokszög területét változói szerint parciálisan deriváljuk és a deriváltakat 0-val tesszük egyenlővé, akkor az adódik, hogy az egyik negyedkörben levő, T-hez legközelebb eső két pont ismeretében a többit a fent leírtakhoz hasonló módon megkaphatjuk.

2) $n = 1, \ldots, 100$ esetre numerikus számolással azt kaptuk, hogy a negyedkör N_n^* által le nem fedett területének n-szerese egységsugarú körnél 0.36 2 tizedesjegy pontosságig.

Köszönetet mondok Ruda Mihálynak, a Valószínűségszámítási osztály munkatársának e cikk megírásához adott hasznos tanácsaiért.

Táblázatok a CDC 3300-as gépen végzett számolás eredményéről:

1. N_n^* pontjainak elhelyezkedése:

n \ ∢	P_0OP_1	P_0OP_2	P_0OP_3	P_0OP_4	P_0OP_5	P_0OP_6	P_0OP_7	P_0OP_8	P_0OP_9	P_0OP_{10}
1	45°									
2	58,31°	31,73°								
3	64,94°	45°	25,1°							
4	68,99°	52,69°	37,35°	21,05°						
5	71,74°	57,78°	45°	32,26°	18,29°					
6	73,76°	61,44°	50,37°	39,64°	28,58°	16,26°				
7	75,31°	64,23°	54,39°	45°	35,67°	25,83°	14,76°			
8	76,53°	66,42°	57,51°	49,13°	40,89°	32,52°	23,61°	13,49°		
9	77,54°	68,21°	60,04°	52,42°	45°	37,63°	30,01°	21,84°	12,51°	
10	78,38°	69,69°	62,13°	55,11°	48,36°	41,70°	34,95°	27,94°	20,37°	11,69°

2. N_n^* területe és a negyedkörből N_n^* által le nem fedett terület n-szerese:

n	N_n^*	$n\left(\frac{\pi}{4} - N_n^*\right)$
1	0.5	0.285398
10	0.748816	0.365822
20	0.767140	0.365163
30	0.773259	0.364160
40	0.776313	0.363418
50	0.778141	0.362869
60	0.779357	0.362447
70	0.780225	0.362115
80	0.780875	0.361844
90	0.781380	0.361619
100	0.781784	0.361430

Summary

An algorithm is given for filling the quadrant of a circle with disjoint parallelograms with sides parallel to to the radii boundaring the quadrant, and with maximal total area.

Резюме

Дается алгоритм для напольнении четверти круга с непересекающимися параллелограмми, со стороному параллельными с граничными прямыми четверти круга и с максимальной общей площадью.

Beérkezett: 1973. szeptember 15.

MTA Számítástechnikai és Automatizálási Kutató Intézete, **Közlemények** 11. (1973)

AZ ÁTLAGKÉSZLET CSÖKKENTÉSE FOLYAMATOS NYILVÁNTARTÁSÚ MODELL BEVEZETÉSÉVEL

dr. Gerencsér László

BEVEZETÉS

Az átlagkészlet csökkentése fontos feladat, amelyet azonban a készletgazdálkodás matematikai elmélete nem vizsgált elégségesen. A közismert készletgazdálkodási modelleken belül az átlagkészlet csökkentése csak az átlaghiány növekedésének az árán biztosítható. Ezzel szemben, ha az egyik modellről egy másik modellre való áttérés realizálható, akkor az előbbi nehézség elkerülhető. A dolgozatban megmutatjuk, hogy ha a S szintre felrendelés elnevezésű készletgazdálkodási modell helyett egy folyamatos nyilvántartású, ún. (r,Q) modellt vezetünk be, akkor az átlagkészlet és az átlaghiány egyszerre csökkenthetők, anélkül, hogy a rendelések átlagos gyakorisága megváltozna. Az így elérhető megtakarításra becsléseket is adunk. Az ilyen vizsgálat további előnye az, hogy elkerüli a gyakran nehezen becsülhető költségtényezők beépítését, a beavatkozás közvetlenül a fizikailag végbemenő folyamatot javítja. Az S-modell illetve az (r,Q) modell részletes ismertetése megtalálható [1]-ben.

1. A MODELLEK LEIRÁSA

Egyetlen cikket készletezünk, amelyet folyamatosan igényelnek a felhasználók. A $(0,t)$ időszakaszban jelentkező igényt $\alpha(t)$ jelöli, ahol $\alpha(t)$ sztochasztikus folyamat. Az $\alpha(t)$ folyamatot Wiener-folyamattal közelítjük, amelynek paraméterei m,σ. Az $\alpha(t)$ változó eloszlásfüggvénye tehát $\Phi\left(\frac{x-mt}{\sigma\sqrt{t}}\right)$, ahol $\Phi(y)$ a standard normális eloszlásfüggvény. A készlet pótlása rendelés útján történik. Egy rendelés megérkezése a rendelés feladását követő τ időpontban történik, ahol τ konstans.

Az S-modellben a felülvizsgálat periódikusan történik. A periódus hossza T, úgy hogy rendeléseket adunk fel a $0, T, 2T\ldots$ időpontokban. A rendelés után a készletállapot (ld. [1]) éppen S. A rendelési mód egy szemléletes megfogalmazása a következő: minden periódusban annyit rendelünk, amennyi a megelőző periódusban elfogyott.

Az (r,Q) modell akkor alkalmazható, amikor a készletállapotban történő minden változást folyamatosan nyilvántartunk. Amikor a készletállapot elér egy kritikus r értéket azonnal feladunk egy rendelést egy Q nagyságú tételre. Igy a készletállapot minden rendelés után ugyanaz az $r+Q$. Világos, hogy az (r,Q) modell érzékenyebb az igényfolyamatban jelentkező ingadozásokra.

Bár a bevezetendő operációk függetlenek a különböző költségektől, ezek a költségek léteznek és hatnak. Bevezetünk háromféle költséget: A, egy rendelés feladásának és elfogadásának költsége, IC egységnyi árunak egységnyi ideig való készletezéséből eredő költség, $\hat{p}$ egységnyi árunak egységnyi ideig fennálló hiányából eredő költség. Az egy időegységre eső átlagos teljes költséget $K_1(S)$, illetve $K_2(r,Q)$ fogják jelölni. Az egy időegységre eső átlagkészletet illetve átlaghiányt D illetve B fogják jelölni, megfelelő alsó index-szel ellátva.

Hogy megkapjuk a $K_1(S)$ költségfüggvény kifejezését először felidézzük a

$$D_1 = B_1 + S - m\tau - \frac{mT}{2} \tag{1.1}$$

formulát (ld. [2]), amelyből könnyen adódik, hogy

$$K_1(S) = \frac{A}{T} + IC\left(S - m\tau - \frac{mT}{2}\right) + (IC + \hat{p})B_1 \tag{1.2}$$

A későbbiek kedvéért részletezzük B_1 kifejezését. Először is

$$B_1 = \frac{1}{T} \int_{\tau}^{\tau+T} B_1(t)dt \tag{1.3}$$

ahol $B_1(t)$ a t időpontban fennálló várható hiány. Ennek kifejezése a következő

$$B_1(t) = \int_S^{\infty} (y - S) \frac{1}{\sigma\sqrt{t}} \varphi\left(\frac{y - mt}{\sigma\sqrt{t}}\right) dy \tag{1.4}$$

Az (r,Q) modell esetén hasonló módon indulhatunk el. Fennáll ugyanis a

$$D_2 = B_2 + \psi + \frac{Q}{2} - m\tau \tag{1.5}$$

egyenlőség. Ebből könnyen adódik a

$$K_2(r,Q) = \frac{mA}{Q} + IC\left(r + \frac{Q}{2} - m\tau\right) + (IC + \hat{p})B_2$$

kifejezés. A B_2 formulája pedig

$$B_2 = \frac{1}{Q} \int_r^{r+Q} B_2(x)dx$$

ahol $B_1(x)$ a τ idő után fennálló várható hiány, ha a kezdőkészletállapot x. $B_2(x)$ integrálalakja a következő:

$$B_2(x) = \int_x^{\infty} (y - x) \frac{1}{\sigma\sqrt{t}} \varphi\left(\frac{y - m\tau}{\sigma\sqrt{\tau}}\right) dy$$

2. A KÉT MODELL MEGFELELTETÉSE

Tegyük fel, hogy egy S modellnek megfelelően történik a készletek utánpótlása. A készletállapot várható értéke közvetlenül a rendelést megelőzően $S - mT$, a rendelt mennyiség várható értéke pedig mT. Amennyiben lehetséges, hogy áttérjünk egy folyamatos nyilvántartású rendszerre, természetes választásként adódnak az

(2.1) $\quad r = S - mT \qquad\qquad Q = mT$

paraméterek. Az előző pont formuláira pillantva látjuk, hogy ez valóban szerencsés választás: vonjuk ki ugyanis (1.5)-ből (1.1)-et, akkor kapjuk, hogy

(2.2) $\quad D_1 - D_2 = B_1 - B_2$

A D és B értékek tehát azonos mennyiséggel változnak meg, ezért elég megmutatni, hogy B_2 kisebb, mint B_1. Vegyük észre azt is, hogy a rendelések átlagos gyakorisága változatlan maradt.

A teljesség kedvéért írjuk fel a teljes átlagos költségek különbségét, melyeket röviden jelöljünk K_1, illetve K_2-vel. Kapjuk, hogy

(2.3) $\quad K_1 - K_2 = (IC + \hat{p})(B_1 - B_2)$

Feladatunk már most a

(2.4) $\quad B_1 \geqslant B_2$

egyenlőtlenség bizonyítása. Ennek céljából a két modell között további megfeleltetést adunk meg. Az egymásnak megfelelő mennyiségek idő illetve készletállapot lesznek. Legyen s egy paraméter, amely 0-tól T-ig változik. Az egymásnak megfelelő értékek legyenek

(2.5) $\quad \tau + s \qquad\qquad S - ms$

A $B_1(t)$ illetve $B_2(x)$ értékeket ezen helyettesítési értékek mellett $\overline{B}_1(s)$ és $\overline{B}_2(s)$ jelölik. A következőkben bebizonyítjuk, hogy

(2.6) $\quad \overline{B}_1(s) \geqslant \overline{B}_2(s)$

3. NORMÁLIS VÁLTOZÓK POZITIV RÉSZÉNEK INTEGRÁLJA

Idézzük fel $\overline{B}_1(s)$ jelentését. Bevezetve a

(3.1) $\quad \xi_1(s) = \alpha(\tau + s) - S$

valószínűségi változót $\overline{B}_1(s)$ nem más, mint e változó pozitív részének a várható értéke, tehát

(3.2) $\quad \overline{B}_1(s) = E(\xi_1^+(s))$

A második modellbe vezessük be a

(3.3) $\quad \xi_2(s) = \alpha(\tau) + ms - S$

valószínűségi változót.

Nyilvánvaló, hogy

(3.4) $\quad \bar{B}_2(s) = E(\xi_2^+(s))$

Vegyük most észre, hogy

(3.5) $\quad E(\xi_1) = E(\xi_2)$

de

(3.6) $\quad D^2(\xi_1) \geqslant D^2(\xi_2)$

a baloldal értéke ugyanis $\sigma^2(\tau + s)$, mígya jobboldalé $\sigma^2\tau$.

Általánosabban, vegyünk egy ξ normális eloszlású valószínűségi változót m,σ paraméterekkel. A mi céljainkhoz elegendő megmutatni, hogy $E(\xi^+)$ σ-nak növekvő függvénye, amikor m fix. Az $E(\xi^+)$ kifejezése a következő alakban írható

(3.7) $$R(m,\sigma) = \int_0^\infty x \frac{1}{\sigma} \varphi\left(\frac{x-m}{\sigma}\right) dx$$

Ez az integrál explicite is kiértékelhető, melynek eredményeképpen

(3.8) $$R(m,\sigma) = \sigma\varphi\left(\frac{m}{\sigma}\right) + m\phi\left(\frac{m}{\sigma}\right)$$

A σ szerinti parciális deriválás eredménye

(3.9) $$\frac{\delta R(m,\sigma)}{\delta\sigma} = \varphi\left(\frac{m}{\sigma}\right)$$

amely állításunkat bizonyítja, hiszen $\varphi(x)$ minden x-re pozitív.

Igy teljes egészében bebizonyítottuk, hogy a folyamatos nyilvántartású rendszer a kívánt tulajdonságokkal bír. Becsléseket is adhatunk figyelembevéve $R(m,\sigma)$-nak σ-ban való konvexitását, amely (3.9)-ből világos. Rutinszámolás után a következő becslést kapjuk

(3.10) $\quad B_1 \geqslant B_2 + \sigma T\gamma$

ahol

(3.11) $$\gamma = \varphi\left(\frac{S - m\tau}{\sigma\sqrt{\tau}}\right)\frac{1}{4\sqrt{\tau + T}}$$

Ez a becslés azon feltétel mellett érvényes, hogy a pillanatnyilag használt S-modell-ben az átlagkészlet nagyobb mint az átlaghiány. Ezt a becslést egyszerűsége miatt érdemes használni, bár finomabb becslések is kaphatók nehézség nélkül.

Irodalom

[1] G. Hadley, T. Whitin, The Analysis of Inventory Systems, Prentice Hall, 1963.

Beérkezett: 1973. szeptember 18.

Summary

Reduction of the one hand inventory by introducing a transaction reporting system

Consider the models (r,Q) and (S,T) described in the book of Hadley and Whitin: Analysis of inventory systems. The average demand is denoted by m its dispersion by 6. Suppose we have an (S,T) model and change it for an (r,Q) model choosing

$$r = S - mT \qquad\qquad Q = mT.$$

We show that with this change both the average an hand inventory and the average stockout is reduced, while the cost of placing orders is unchanged. The reduction is approximately proportional with σ.

Р е з ю м е

УМЕНЬШЕНИЕ СРЕДНИЙ УРОВЕНЬ ЗАПАСОВ ПРИ ИСПОЛЬЗОВАНИИ ПОДЕЛЕМ С НЕПРЕРЫВНЫМ НАБЛЮДЕНИЕМ

Рассматриваются модели (r, Q) и (S, T) описаны книге Хедли и Уайтина. Средний спрос обозначим через m, его дисперсию через 6. Перейдём от модели (S, T) К модели (r, Q) с помощью уравнений

$$r = S - mT \qquad\qquad Q = mT$$

Покажем что при переходе как и средний уровень запасов так и средний уровень невыполнянных спросов уменьшается, но стоимость заказов не изменяется. Уменьшение приблизительно имейно зависит от σ.

MTA Számítástechnikai és Automatizálási Kutató Intézete, **Közlemények** 11. (1973)

GAUSS FOLYAMATOK ELTOLÁSI PARAMÉTERÉNEK BAYES FÉLE BECSLÉSE

Pham Ngoc Phuc

1. Tekintsük az

(1) $$y(j) = \xi(j) + \Theta, \qquad j = 1, 2, \dots, n$$

sztochasztikus folyamatot.

Tegyük fel, hogy $\xi(j)$ a következő elsőrendű autoregressziós folyamat

(2) $$\xi(1) = \epsilon(1)$$
$$\xi(k) = \lambda\xi(k-1) + \epsilon(k), \qquad k = 2, \dots, n$$

ahol $\epsilon(j)$ független sorozat, $F_j(x)$ eloszlással. A Θ $(\Theta \in R^1)$ paraméter becslésének Bayes-féle megfogalmazásában feltesszük, hogy Θ olyan valószínűségi változó, amelynek (a priori) eloszlása $\Pi(\Theta)$, továbbá Θ független $\epsilon(1), \dots, \epsilon(n)$-től.

A sztochasztikus folyamatok elméletéből (lásd pl. [2]) tudjuk, hogy Θ legjobb (torzítatlan) becslése a

$$\hat{\Theta} = E(\Theta \mid y(1), \dots, y(n))$$

statisztika*, és Θ legjobb lineáris (torzítatlan) becslése

$$\hat{\Theta}_1 = \hat{E}(\Theta \mid H_y),$$

ahol H_y jelöli az $y(1), \dots, y(n)$ változók által generált zárt lineáris sokaságot, és $\hat{E}(\cdot \mid H_y)$ a H_y-ra való vetületét jelenti.

Abban az esetben, amikor a $(\Theta, y(1), \dots, y(n))$ valószínűségi vektorváltozó Gauss eloszlású, akkor Θ legjobb becslése megegyezik Θ legjobb lineáris becslésével, azaz

(3) $$E(\Theta \mid y(1), \dots, y(n)) = \hat{E}(\Theta \mid H_y).$$

Vizsgáljuk ennek a megfordítását. Megmutatjuk, hogy az (1)-(2) séma esetén ha (3) teljesül, akkor bizonyos egyszerű feltétel mellett, $\Pi(\Theta), F_j(x)$ Gauss-eloszlások.

A bizonyítás Kagan-Karpov [1] cikk gondolatmenetén alapszik.

* Ebben a dolgozatban a becslések optimalitásának fogalma a négyzetes veszteségfüggvényhez kapcsolódik.

Bevezetjük az

$$
(4)\qquad
\begin{aligned}
x(1) &= \epsilon(1) + \Theta \\
x(2) &= \epsilon(2) + \Theta \\
&\ldots\ldots\ldots\ldots \\
x(n) &= \epsilon(n) + \Theta
\end{aligned}
$$

változókat.

Belátható, hogy

$$
(5)\qquad
\begin{aligned}
y(1) &= x(1) \\
y(2) &= x(2) + \lambda\epsilon(1) \\
&\ldots\ldots\ldots\ldots\ldots \\
y(n) &= x(n) + \lambda^{n-1}\epsilon(1) + \ldots + \lambda\epsilon(n-1).
\end{aligned}
$$

H_x-szel jelöljük az $x(1), \ldots, x(n)$ változók által generált zárt lineáris sokaságot. (5)-ből következik, hogy H_y megegyezik H_x-szel, amiből

$$
(6)\qquad \hat{E}(\Theta \mid H_y) = \hat{E}(\Theta \mid H_x)
$$

adódik.

Legyenek

$$
\begin{aligned}
&E\epsilon(j) = \mu_1^{(j)}, && E\Theta = \alpha_1 \\
&E(\epsilon(j) - \mu_1^{(j)})^k = \mu_k^{(j)}, && E(\Theta - \alpha_1)^k = \alpha_k \qquad k = 2, 3, \ldots,
\end{aligned}
$$

Tegyük fel, hogy

$$
(7)\qquad 0 < \alpha_2 < \infty\,; \qquad 0 < \mu_2^{(j)} < \infty, \qquad j = 1, 2, \ldots, n.
$$

Ismeretes (lásd [1]), hogy az $x(j) = \epsilon(j) + \Theta$ séma esetén Θ legjobb lineáris becslése a tágabb értelemben vett Bayes-féle becslés és a következő alakú:

$$
(8)\qquad \hat{E}(\Theta \mid H_x) = c_0 + \sum_{j=1}^{n} c_j x(j)
$$

ahol

$$
(9)\qquad c_0 = \frac{\alpha_1 - \sum\limits_{j=1}^{n} \dfrac{\alpha_2 \mu_1^{(j)}}{\mu_2^{(j)}}}{1 + \sum\limits_{j=1}^{n} \dfrac{\alpha_2}{\mu_2^{(j)}}}, \qquad c_j = \frac{\dfrac{\alpha_2}{\mu_2^{(j)}}}{1 + \sum\limits_{j=1}^{n} \dfrac{\alpha_2}{\mu_2^{(j)}}} \qquad j = 1, 2, \ldots, n.
$$

Ily módon (6), (8)-ból kapjuk, hogy az $y(j) = \xi(j) + \Theta$ folyamat esetén, ahol $\xi(j)$ (2) alakú elsőrendű autoregressziós folyamat, és $\epsilon(j)$-re, Θ-ra teljesül a (7) feltétel, Θ legjobb lineáris becslése

$$(10) \qquad \hat{\Theta}_1 = \hat{E}(\Theta \mid H_y) = c_0 + \sum_{j=1}^{n} c_j y(j) - (\lambda c_2 + \ldots + \lambda^{n-1} c_n)\epsilon_1 - \ldots - \lambda c_n \epsilon_{n-1} ,$$

ahol c_0, c_j, $j = 1, \ldots, n$ együtthatók meghatározhatók (9) segítségével.

Legyenek

$$f_j(t) = \int_{-\infty}^{+\infty} e^{itx} dF_j(x) , \qquad g_j(t) = \frac{f_j'(t)}{f_j(t)} \qquad j = 1, 2, \ldots, n.$$

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{it\Theta} d\Pi(\Theta) , \qquad \psi(t) = \frac{\varphi'(t)}{\varphi(t)} ,$$

ahol a $g_j(t)$, $\psi(t)$ függvények a $t = 0$ pont valamilyen környezetében vannak értelmezve.

Tegyük most fel, hogy Θ legjobb becslése megegyezik Θ legjobb lineáris becslésével, így (3)-ból és (10)-ből következik, hogy

$$(11) \qquad E(\Theta \mid y_1, \ldots, y_n) = c_0 + \sum_{j=1}^{n} c_j y_j - (\lambda c_2 + \ldots + \lambda^{n-1} c_n)\epsilon_1 - \ldots - \lambda c_n \epsilon_{n-1}$$

Az általánosság korlátozása nélkül feltehetjük, hogy $\alpha_1 = \mu_1^{(j)} = 0$ $j = 1, \ldots, n$; tehát $c_0 = 0$.

Legyenek

$$(12) \qquad \begin{aligned} -(\lambda c_2 + \ldots + \lambda^{n-1} c_n) &= r_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ -\lambda c_n &= r_{n-1} \end{aligned}$$

akkor (11) az

$$(13) \qquad E(\Theta \mid y_1, \ldots, y_n) = \sum_{j=1}^{n} c_j y_j + r_1 \epsilon_1 + \ldots + r_{n-1} \epsilon_{n-1} = \sum_{j=1}^{n} c_j y_j + \sum_{j=1}^{n-1} r_j \epsilon_j$$

alakra hozható. A feltételes várható érték jól ismert tulajdonsága és (13) alapján kapjuk az

$$(14) \qquad E\left(\Theta \exp i \sum_{j=1}^{n} t_j y_j\right) = E\left[\left(\exp i \sum_{j=1}^{n} t_j y_j\right) E(\Theta \mid y_1, \ldots, y_n)\right] = E\left(\sum_{j=1}^{n} c_j y_j \exp i \sum_{j=1}^{n} t_j y_j\right) +$$
$$+ E\left(\sum_{j=1}^{n-1} r_j \epsilon_j \exp i \sum_{j=1}^{n} t_j y_j\right)$$

összefüggést.

Mivel

$$(15)\qquad \sum_{j=1}^{n} t_j y_j = \sum_{j=1}^{n} t_j(\lambda^{j-1}\epsilon_1 + \ldots + \lambda\epsilon_{j-1} + \epsilon_j + \Theta) = \nu_1\epsilon_1 + \nu_2\epsilon_2 + \ldots + \nu_n\epsilon_n + \tau\Theta =$$

$$= \tau\Theta + \sum_{j=1}^{n} \nu_j\epsilon_j$$

ahol

$$(16)\qquad \begin{aligned} \nu_1 &= t_1 + \lambda t_2 + \ldots + \lambda^{n-1} t_n \\ \nu_2 &= t_2 + \lambda t_3 + \ldots + \lambda^{n-2} t_n \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ \nu_n &= t_n \end{aligned}$$

és

$$\tau = \sum_{j=1}^{n} t_j = \nu_1 + (1-\lambda)\sum_{j=2}^{n} \nu_j ,$$

látható, hogy

$$(17)\qquad E\left(\Theta \exp i \sum_{j=1}^{n} t_j y_j\right) = E\left(\Theta \cdot e^{i\tau\Theta + i\sum_{j=1}^{n}\nu_j\epsilon_j}\right) = E(\Theta \cdot e^{i\tau\Theta})E\left(e^{i\sum_{j=1}^{n}\nu_j\epsilon_j}\right) = -i\varphi'(\tau)\prod_{j=1}^{n} f_j(\nu_j)$$

Tekintsük (14) jobb oldalát.

Egyrészt

$$(18)\qquad E\left(\sum_{j=1}^{n} c_j y_j \exp i \sum_{j=1}^{n} t_j y_j\right) = E\left(\sum_{j=1}^{n} u_j\epsilon_j + \gamma\Theta\right)\exp i\left(\sum_{j=1}^{n} \nu_j\epsilon_j + \tau\Theta\right)$$

$$= E\left(\sum_{j=1}^{n} u_j\epsilon_j e^{i\sum_{j=1}^{n}\nu_j\epsilon_j + i\tau\Theta} + \gamma\Theta e^{i\sum_{j=1}^{n}\nu_j\epsilon_j + i\tau\Theta}\right)$$

$$= E\left(\sum_{j=1}^{n} u_j\epsilon_j e^{i\sum_{j=1}^{n}\nu_j\epsilon_j}\right) E(e^{i\tau\Theta}) + E(\gamma\Theta e^{i\tau\Theta})E\left(e^{i\sum_{j=1}^{n}\nu_j\epsilon_j}\right)$$

$$= -i\sum_{j=1}^{n} u_j f_j'(\nu_j)\prod_{k\neq j} f_k(\nu_k)\varphi(\tau) - i\gamma\varphi'(\tau)\prod_{j=1}^{n} f_j(\nu_j) ,$$

ahol

ahol

$$(19)\qquad \begin{aligned} u_1 &= c_1 + \lambda c_2 + \ldots + \lambda^{n-1} c_n \\ u_2 &= c_2 + \lambda c_3 + \ldots + \lambda^{n-2} c_n \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ u_n &= c_n \end{aligned}$$

és

$$\gamma = c_1 + c_2 + \ldots + c_n\,.$$

Másrészt

$$E\left(\sum_{j=1}^{n-1} r_j\epsilon_j \exp i \sum_{j=1}^{n} t_j y_j\right) = E\left(\sum_{j=1}^{n-1} r_j\epsilon_j e^{i(\sum_{j=1}^{n} \nu_j\epsilon_j + \tau\Theta)}\right) = E\left(\sum_{j=1}^{n-1} r_j\epsilon_j e^{i\sum_{j=1}^{n}\nu_j\epsilon_j}\right)\cdot E(e^{i\tau\Theta})$$

$$= E\left(\sum_{j=1}^{n-1} r_j\epsilon_j e^{i\sum_{j=1}^{n-1}\nu_j\epsilon_j}\right) E(e^{i\nu_n\epsilon_n}) E(e^{i\tau\Theta})$$

$$= -i\sum_{j=1}^{n-1} r_j f_j'(\nu_j) \prod_{\substack{k=1\\(k\neq j)}}^{n-1} f_k(\nu_k)\cdot f_n(\nu_n)\cdot\varphi(\tau)$$

Ily módon (14), (17), (18) és (20) alapján kapjuk a következő relációt:

$$(21)\qquad \varphi'(\tau)\prod_{j=1}^{n} f_j(\nu_j) = \varphi(\tau)\sum_{j=1}^{n} u_j f_j'(\nu_j)\prod_{k\neq j} f_k(\nu_k) + \gamma\varphi'(\tau)\prod_{j=1}^{n} f_j(\nu_j) + \varphi(\tau) f_n(\nu_n)\sum_{j=1}^{n-1} r_j f_j'(\nu_j)\cdot \prod_{\substack{k=1\\k\neq j}}^{n-1} f_k(\nu_k)$$

Mivel $\varphi(\tau)\prod_{j=1}^{n} f_j(\nu_j)$ folytonos a $(\nu_1,\ldots,\nu_n) = (0,\ldots,0)$ pontban és $\varphi(0)\prod_{j=1}^{n} f_j(0) = 1$, létezik olyan ϵ-környezet, amelyben $\varphi(\tau)\prod_{j=1}^{n} f_j(\nu_j)$ különbözik 0-tól. (21) mindkét oldalát $\varphi(\tau)\prod_{j=1}^{n} f_j(\nu_j)$-val osztva kapjuk az

$$(1-\gamma)\,\psi(\tau) = \sum_{j=1}^{n} u_j g_j(\nu_j) + \sum_{j=1}^{n-1} r_j g_j(\nu_j)$$

összefüggést, amiből $u_j + r_j = c_j$, $j = 1,\ldots,n-1$; $u_n = c_n$ miatt

$$(22)\qquad (1-\gamma)\,\psi(\tau) = \sum_{j=1}^{n} c_j g_j(\nu_j)$$

adódik. A (7) feltétel teljesülése esetén $\psi(t), g_j(t)$ differenciálható és

$$c_j > 0\,, \qquad 1-\gamma = \frac{1}{1+\sum\limits_{j=1}^{n} \dfrac{\alpha_2}{\mu_2^{(j)}}} > 0\,.$$

(22) mindkét oldalát ν_1, majd ν_j $(j = 2, \ldots, n)$ szerint differenciálva

(23a) $\quad (1-\gamma)\psi'(\tau) = c_1 g_1'(\nu_1)\,,$

(23b) $\quad (1-\gamma)\psi'(\tau)(1-\lambda) = c_j g_j'(\nu_j)$-hez

jutunk. Legyen $\nu_1 = 0$, (23a)-ból következik

(24) $\quad \psi'(t) = \mathrm{cons}\,t = -s^2 < 0$

a $t = 0$ valamilyen környezetében.

Feltesszük, hogy $1-\lambda > 0$; (23a), (23b), (24)-ből belátható, hogy

(25) $\quad g_j'(t) = \mathrm{cons}\,t = -s_j^2 < 0 \qquad j = 1, 2, \ldots, n$

ha $|t|$ eléggé kicsiny.

Mivel $\varphi(0) = f_j(0) = 1$ és $\varphi'(0) = f_j'(0) = 0$, (24), (25) alapján

$$\varphi(t) = e^{-\frac{s^2}{2}t^2}$$

$$f_j(t) = e^{-\frac{s_j^2}{2}t^2}$$

a $t = 0$ egy környezetében, ahonnan adódik, hogy $\Pi(\Theta)$, $F_j(x)$ Gauss eloszlású. Ily módon igazoltuk a következő állítást:

1. Tétel. *Tegyük fel, hogy az* (1)-(2) *sémában* $n \geqslant 2$, $1-\lambda > 0$ *és teljesül a* (7) *feltétel.*

A Θ paraméter legjobb becslése akkor és csak akkor egyezik meg a legjobb lineáris becsléssel, ha $\Pi(\Theta)$ normális eloszlás és $y(j)$ Gauss folyamat.

Megjegyzés. Megmutatható, hogy a fenti állítás – bizonyos egyszerű feltételek teljesülése esetén – érvényben marad abban az esetben, amikor $\xi(j)$ a következő p-edrendű autoregressziós folyamat:

$$\xi(1) = \epsilon(1)$$
$$\xi(2) + a_1 \xi(1) = \epsilon(2)$$
$$\dots\dots\dots\dots\dots\dots$$
$$\xi(p) + a_1 \xi(p-1) + \dots + a_{p-1} \xi(1) = \epsilon(p)$$
$$\xi(k) + a_1 \xi(k-1) + \dots + a_p \xi(k-p) = \epsilon(k) \qquad k = p+1, \dots, n\,,$$

ahol $\epsilon(j)$, $j = 1, 2, \dots, n$ független sorozat, $F_j(x)$ eloszlással és $(\epsilon_1, \dots, \epsilon_n)$ független Θ-tól.

2. Tekintsük most az (1)-(2) sémát azzal a feltevéssel, hogy $\epsilon(j)$, $j = 1, \dots, n$ független sorozat, azonos $F(x)$ aloszlással.
Ebben az esetben az

$$E\epsilon(j) = \mu_1\,, \qquad E\Theta = \alpha_1\,,$$
$$E[\epsilon(j) - \mu_1]^k = \mu_k\,, \qquad E(\Theta - \alpha_1)^k = \alpha_k\,, \qquad k \geqslant 2$$

jelölésekkel belátható, hogy a

(26) $$0 < \alpha_2 < \infty$$
$$0 < \mu_2 < \infty$$

feltétel teljesülése esetén Θ-nak (10) legjobb lineáris becslése a következő alakú:

(27) $$\hat{\Theta}_1 = \hat{E}(\Theta \mid H_y) = c_0 + c \sum_{j=1}^{n} y(j) + c \sum_{j=1}^{n-1} \rho_j \epsilon(j)$$

ahol

$$c_0 = \frac{\alpha_1 - n\alpha_2 \dfrac{\mu_1}{\mu_2}}{1 + n\dfrac{\alpha_2}{\mu_2}}\,, \qquad c = \frac{\dfrac{\alpha_2}{\mu_2}}{1 + n\dfrac{\alpha_2}{\mu_2}} = \frac{\alpha_2}{\mu_2 + n\alpha_2}\,,$$

és

$$\rho_1 = -(\lambda + \dots + \lambda^{n-1})$$
$$\dots\dots\dots\dots\dots\dots$$
$$\rho_{n-1} = -\lambda$$

Legyen $l = \sum_{j=1}^{n} b_j y(j)$ egy általános alakú lineáris becslés. Vizsgáljuk azt a problémát, hogy milyen esetben áll fenn a következő reláció:

$$(28)\qquad E\left(\Theta \mid \sum_{j=1}^{n} y(j) + \sum_{j=1}^{n-1} \rho_j \epsilon(j),\ \sum_{j=1}^{n} b_j y(j)\right) = \hat{\Theta}_1 = c_0 + c\left(\sum_{j=1}^{n} y(j) + \sum_{j=1}^{n-1} \rho_j \epsilon(j)\right).$$

Az általánosság korlátozása nélkül feltehetjük, hogy $\alpha_1 = \mu_1 = 0$, ekkor $c_0 = 0$ és a (28) formula az

$$(29)\qquad E\left(\Theta \mid \sum_{j=1}^{n} y(j) + \sum_{j=1}^{n-1} \rho_j \epsilon(j),\ \sum_{j=1}^{n} b_j y(j)\right) = c\sum_{j=1}^{n} y(j) + c\sum_{j=1}^{n-1} \rho_j \epsilon(j)$$

alakra hozható.

Tegyük most fel, hogy (29) teljesül. Ebben az esetben

$$(30)\qquad E\left[\Theta \exp i\left(t\sum_{j=1}^{n} y_j + t\sum_{j=1}^{n-1} \rho_j \epsilon_j + z\sum_{j=1}^{n} b_j y_j\right)\right] = E\left[\exp i\left(t\sum_{j=1}^{n} y_j + t\sum_{j=1}^{n-1} \rho_j \epsilon_j + \right.\right.$$
$$\left.\left. + z\sum_{j=1}^{n} b_j y_j\right) E\left(\Theta \mid \sum_{j=1}^{n} y_j + \sum_{j=1}^{n-1} \rho_j \epsilon_j, l\right)\right]$$
$$= E\left[c\left(\sum_{j=1}^{n} y_j + \sum_{j=1}^{n-1} \rho_j \epsilon_j\right) \exp i\left(t\sum_{j=1}^{n} y_j + \right.\right.$$
$$\left.\left. + t\sum_{j=1}^{n-1} \rho_j \epsilon_j + z\sum_{j=1}^{n} b_j y_j\right)\right]$$

Mivel

$$\sum_{j=1}^{n} y_j + \sum_{j=1}^{n-1} \rho_j \epsilon_j = n\Theta + \sum_{j=1}^{n} \epsilon_j,$$

$$\sum_{j=1}^{n} b_j y_j = \beta_1 \epsilon_1 + \ldots + \beta_n \epsilon_n + (b_1 + \ldots + b_n)\Theta = \sum_{j=1}^{n} \beta_j \epsilon_j + n\beta\Theta,$$

ahol

$$\beta_1 = b_1 + \lambda b_2 + \ldots + \lambda^{n-1} b_n$$
$$\beta_2 = b_2 + \lambda b_3 + \ldots + \lambda^{n-2} b_n$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\beta_n = b_n,$$

és

$$\beta = \frac{1}{n}(b_1 + b_2 + \ldots + b_n),$$

(30)-ból kapjuk az

$$(31) \qquad E\left\{\Theta \exp i\left[t\left(n\Theta + \sum_{j=1}^{n} \epsilon_j\right) + z\left(n\beta\Theta + \sum_{j=1}^{n} \beta_j \epsilon_j\right)\right]\right\} = E\left\{c\left(n\Theta + \sum_{j=1}^{n} \epsilon_j\right) \exp i\left[t\left(n\Theta + \sum_{j=1}^{n} \epsilon_j\right) + z\left(n\beta\Theta + \sum_{j=1}^{n} \beta_j \epsilon_j\right)\right]\right\}$$

összefüggést.

Belátható, hogy egyrészt

$$(32) \qquad E\left\{\Theta \exp i\left[t\left(n\Theta + \sum_{j=1}^{n} \epsilon_j\right) + z\left(n\beta\Theta + \sum_{j=1}^{n} \beta_j \epsilon_j\right)\right]\right\} = E\left[\Theta \cdot e^{i(tn + zn\beta)\Theta + i\sum_{j=1}^{n}(t + z\beta_j)\epsilon_j}\right]$$

$$= E\left[\Theta \cdot e^{i(tn + zn\beta)\Theta}\right] \cdot E\left[e^{i\sum_{j=1}^{n}(t + z\beta_j)\epsilon_j}\right]$$

$$= -i\varphi'(nt + n\beta z)\prod_{j=1}^{n} f(t + z\beta_j),$$

ahol $f(t) = \int_{-\infty}^{+\infty} e^{itx}\, dF(x)$, $\varphi(t) = \int_{-\infty}^{+\infty} e^{it\Theta}\, d\Pi(\Theta)$, másrészt

$$(33) \qquad E\left[c\left(n\Theta + \sum_{j=1}^{n} \epsilon_j\right) e^{i(nt + n\beta z)\Theta + i\sum_{j=1}^{n}(t + z\beta_j)\epsilon_j}\right] =$$

$$= E\left[cn\Theta e^{i(nt + n\beta z)\Theta}\right] E\left[e^{i\sum_{j=1}^{n}(t + z\beta_j)\epsilon_j}\right] + E\left[c\sum_{j=1}^{n} \epsilon_j e^{i\sum_{j=1}^{n}(t + z\beta_j)\epsilon_j}\right] \cdot E\left[e^{i(nt + n\beta z)\Theta}\right]$$

$$= -icn\varphi'(nt + n\beta z)\prod_{j=1}^{n} f(t + z\beta_j) - ic \cdot \varphi(nt + n\beta z)\sum_{j=1}^{n} f'(t + \beta_j z)\prod_{k \neq j} f(t + \beta_k z)$$

Ily módon (31), (32) és (33)-ból következik

$$-i\varphi'(nt + n\beta z)\prod_{j=1}^{n} f(t + z\beta_j) = -icn\varphi'(nt + n\beta z)\prod_{j=1}^{n} f(t + z\beta_j) - ic\varphi(nt + n\beta z)\sum_{j=1}^{n} f'(t + \beta_j z)\prod_{k \neq j} f(t + \beta_k z),$$

ahonnan

$$(34) \qquad \frac{\varphi'(nt + n\beta z)}{\varphi(nt + n\beta z)}(1 - cn) = c \sum_{j=1}^{n}{}' \frac{f'(t + \beta_j z)}{f(t + \beta_j z)}$$

adódik, ha $|t|$ és $|z|$ eléggé kicsiny.

$\psi(t) = \frac{\varphi'(t)}{\varphi(t)}$, $g(t) = \frac{f'(t)}{f(t)}$ jelölésekkel, $\frac{c}{1 - nc} = \frac{\alpha_2}{\mu_2} > 0$ miatt, (34)-ből kapjuk a

$$(35) \qquad \psi(nt + n\beta z) = \frac{c}{1 - nc} \sum_{j=1}^{n}{}' g(t + \beta_j z)$$

összefüggést.

Legyen $t = -\beta z$, így (35)-ből

$$(36) \qquad \sum_{j=1}^{n}{}' g(\delta_j z) = 0$$

adódik, ahol $\delta_j = \beta_j - \beta$, $|z| < \epsilon$, ϵ eléggé kicsiny.

Tegyük fel, hogy teljesülnek a következő feltételek:

(37) δ_j, $j = 1, \ldots, n$,

közül létezik pontosan egy δ_k abszolút értékben legnagyobb, mondjuk δ_n, azaz

$$|\delta_n| > \max \; |\delta_1|, \ldots, |\delta_{n-1}| \; ,$$

továbbá

$$\frac{\delta_j}{\delta_n} < 0 \quad j = 1, \ldots, n - 1.$$

(38) $g(t)$ deriváltja az origóban folytonos (vagy ami ezzel ekvivalens, $f''(t)$ az origóban folytonos).

[3]-ban bebizonyították, hogy (37), (38) teljesülése esetén a (36) egyenletből következik, hogy $F(x)$ normális eloszlásfüggvény. Ilyen esetben (35)-ből belátható, hogy

$$\psi(t) = -q^2 t$$

alakú (q állandó), ha $|t|$ eléggé kicsiny, ami azt jelenti, hogy $\Pi(\Theta)$ normális eloszlás.

Megfordítva, $y(j)$ Gauss folyamat és $\Pi(\Theta)$ normális eloszlás esetén megmutattuk, hogy $E(\Theta | y_1, \ldots, y_n) = c\left(\sum_{j=1}^{n}{}' y_j + \sum_{j=1}^{n-1}{}' \rho_j \epsilon_j\right)$ és még inkább $E\left(\Theta | \sum_{j=1}^{n}{}' y_j + \sum_{j=1}^{n-1}{}' \rho_j \epsilon_j, l\right) =$

$$= c\left(\sum_{j=1}^{n}{}' y_j + \sum_{j=1}^{n-1}{}' \rho_j \epsilon_j\right).$$

Tehát igazoltuk a következő állítást:

2. Tétel. *Tegyük fel, hogy* $n \geqslant 2$ *és teljesülnek a (26), (37), (38) feltételek. Az*

$$E\left(\Theta \mid \sum_{j=1}^{n} y_j + \sum_{j=1}^{n-1} \rho_j \epsilon_j \,,\; \sum_{j=1}^{n} b_j y_j\right) = c_0 + c\left(\sum_{j=1}^{n} y_j + \sum_{j=1}^{n-1} \rho_j \epsilon_j\right)$$

reláció akkor és csak akkor áll fenn, ha $y(j)$ *Gauss-folyamat és* $\Pi(\Theta)$ *normális eloszlás.*

Hálámat és köszönetemet szeretném ezúton is kifejezni Arató Mátyás professzornak a sokoldalú és nagyon értékes tanácsaiért és segítségéért.

Irodalom

[1] А.М. Каган – Ю.Н. Карпов: Байесовская постановка задачи оценивания параметра сдвига. Зап.науч. семинаров ленингр. отд.мат. ин-та АН СССР 1972, 29, 62-73.

[2] И.И. Гихман – А.В. Скороход: Введение в теорию случайных процессов, М.1965.

[3] Rao C.R.: On some characterizations of the normal law. Sankhyā, Ser. A, 29, 1(1967) 1-14.

Beérkezett: 1973. szeptember 19.

Summary

ON THE BAYES' ESTIMATOR OF LOCATION PARAMETER IN THE CASE OF GAUSSIAN PROCESS

In this paper we consider the stochastic process:

(1) $y(j) = \xi(j) + \Theta \qquad j = 1, \ldots, n,$

where $\xi(j)$ is the first order autoregressive process, and Θ is a random variable with distribution $\Pi(\Theta)$.

It is given some conditions of linearity of the Bayes' estimate of the location parameter in the case of the scheme (1).

In particular, Theorem 1 shows that with certain simple assumptions the best estimate and the best linear estimate of Θ are equivalent if and only if $\Pi(\Theta)$ is normal and the process $y(j)$ is Gaussian.

Р е з ю м е

О БАЙЕСОВСКОЙ ПОСТАНОВКЕ ЗАДАЧИ ОЦЕНИВАНИЯ ПАРАМЕТРА СДВИГА В СЛУЧАЕ ГАУССОВКИХ ПРОЦЕССОВ

В этой работе рассматривается случайный процесс:

(1) $y(j) = \xi(j) + \Theta \qquad j = 1, \ldots, n,$

где $\xi(j)$ авторегрессионный процесс первого порядка, и Θ случайная величина с распределением $\Pi(\Theta)$.

Указаны некоторые условия линейности байесовской оценки параметра сдвига Θ в случае схемы (1).

В частности, в теореме 1 показывается, что при некоторых простых условия наилушая оценка и наилушая линейная оценка для Θ эквиваленты тогда и только тогда, когда $\Pi(\Theta)$ нормальное распределение и $y(j)$ гауссовский процесс.

MTA Számítástechnikai és Automatizálási Kutató Intézete, **Közlemények** 11. (1973)

ELSŐRENDŰ AUTÓREGRESSZIÓS FOLYAMAT PARAMÉTERBECSLÉSÉRŐL

Ruda Mihály

J.E. REEVES [6] dolgozatában a

$$(1) \qquad \lambda = \sum_{i=2}^{N} Y_i Y_{i-1} / \sum_{i=2}^{N} Y_{i-1}^2$$

becslés eloszlását vizsgálja az

$$(2) \qquad Y_i = \lambda Y_{i-1} + E_i \,, \qquad i = 2, 3, \ldots, N \,,$$

folyamat Y_i, $i = 1, 2, \ldots, N$ értékeinek ismeretében, ahol Y_1 Gauss eloszlású változó, λ tetszőleges valós szám, E_i független normális sorozat.

Speciális eredmények régebbről ismertek: $\lambda = 0$ értékre ANDERSON [1] adta meg $\hat{\lambda}$ eloszlását. Az N-nel asszimptotikusan végtelenhez tartva $|\lambda| < 1$ értékekre MANN és WALD [5] szerint $\hat{\lambda}$ eloszlása $N(\lambda, (1 - \lambda^2)/N)$, a $|\lambda| > 1$ értékekre viszont WHITE [8] szerint $\hat{\lambda}$ eloszlása olyan, hogy $|\lambda|^N (\hat{\lambda} - \lambda)/(\lambda^2 - 1)$ Cauchy eloszlású.

J.E. REEVES dolgozatában kis elemszámú megfigyelés mellett kívánja a $\hat{\lambda}$ becslés eloszlását meghatározni tetszőleges valós λ érték mellett. Igy többek között a (2) folyamat stacionaritásának vizsgálata is lehetővé válik.

Az eloszlásfüggvény értékét az $x = (1 + \lambda^2)/2\lambda$ pontban pontosan meghatározza, más helyeken azonban csak közelítő megoldást ad χ^2 változók lineáris kombinációinak becslésével, felhasználva SATTERTHWAITE [7] és IMHOF [4] eredményeit.

Ezekre a bonyolult számításokra a $|\lambda| < 1$ (stacionárius) esetben azonban nincs szükség, mivel a λ egy más becslésének eloszlása már táblázatos formában rendelkezésre áll; ld. ARATÓ és BENCZÚR [2], [3]. Ezekben a dolgozatokban ([2], [3]) a paraméterbecslés eloszlásának meghatározása nem a diszkrét folyamat vizsgálatával, hanem a

$$d\xi(t) = -\rho\xi(t)dt + dw(t)$$

sztochasztikus differenciálegyenletnek eleget tevő stacionárius $(\rho > 0)$, Gauss-Markov folyamat ρ paraméterbecslés eloszlásának meghatározásával történik, ahol $0 \leqslant t \leqslant T$, $W(t)$ standard Wiener folyamat: $Mdw = 0$, $M(dw)^2 = \sigma_w^2 dt$, tehát $M\xi(t) = 0$, $M\xi(S)\xi(S + t) = \sigma_\xi^2 e^{-\rho|t|}$ $\sigma_\xi^2 = \sigma_w^2 / 2\rho$).

A ρ paraméter maximum likelihood becslésének pontos eloszlását (a ρ függvényben) ARATÓ és BENCZÚR [2] adja meg (a függvényértékek táblázatával együtt). A diszkrét folyamatot a $0 \leqslant t \leqslant T$ intervallumon megfigyelt folytonos folyamat $i \cdot \delta$ $(i = 0, 1, \ldots, N - 1,$ $(N - 1) \cdot \delta = T)$ pontokban felvett értékeinek tekintve, a λ paramétert a

$$(3) \qquad \lambda = e^{-\rho \cdot \delta} \qquad (\text{ha } T = 1, \ \lambda = e^{-\rho/N})$$

összefüggésből nyerhetjük.

Mivel diszkrét esetben csak véges sok megfigyelésünk van, és nem egy folytonos realizációból becsüljük a λ paramétert, ezért a folytonos folyamat esetének megfelelő pontos eloszlás diszkrét sorozatra nem adható meg. A folytonos folyamatra kapott becslés alapján azonban a diszkrét folyamat paraméterbecslésének eloszlását is elég jó közelítéssel megkaphatjuk, már 10-es nagyságrendű megfigyelésszám mellett is; ld. [3].

A [2] dolgozatban szereplő táblázat $T = 1$-re adja meg a ρ maximum likelihood becslésének eloszlását. Ebből a táblázatból éppen a (3) összefüggés alapján kapunk a λ megfelelő becslésének eloszlására egy jó közelítést. Mivel mind a folytonos, mind a diszkrét esetben eddig ez az egyetlen analitikus úton meghatározott és táblázatba foglalt eloszlás, ezért a gyakorlatban használt becsléseknek a maximum likelihood becsléssel való összehasonlítása és az időben folytonos illetve időben diszkrét modell alapján nyert eredmények összevetése igen fontos feladat. Ezzel foglalkozik a [3] dolgozat, amikor CDC 3300-as gépen futó programok szimulációs eredményeit vizsgálja.

Hasonlítsuk össze a [6]-ban a $|\lambda| < 1$ esetben az (1) becslés eloszlására adott táblázatos értékeket, a folytonos közelítésből (a [2] dolgozat táblázatából) adódó becslésekkel – illetve a normális eloszlással való közelítésből nyert becslésekkel!

A $\lambda = 0.5$ és $\lambda = 0.9$ értékekre az $x = (1 + \lambda^2)/2\lambda$ helyen a pontos függvényérték szerepel REEVES [6] 2. táblázatában. Ezek az értékek azonos x értékekre nagyobbak, mint a normális közelítésből adódó valószínűségek, de a maximum likelihood becslésből kapott határok (azonos valószínűségek mellett) lényegesen szűkebbek (ld. a mellékelt táblázatot). Összehasonlításul megadjuk még a [3] 1. táblázatból vett szimulációs eredményeket az (1) becslésre és a maximum likelihood becslésre is.

Táblázat: $p = p_\lambda(\hat{\lambda} < x)$

λ	normális közelítés, megfigyelésszám 10	normális közelítés, megfigyelésszám 20	az (1) becslés, megfigyelésszám 11	az (1) becslés, megfigyelésszám 21	a folytonos közelítésből adódó becslés, megfigyelésszám 10	a folytonos közelítésből adódó becslés, megfigyelésszám 20	szimulációs eredmények, (1) becslés, megfigyelésszám 20	szimulációs eredmények, max.lik. becslés, megfigyelésszám 20
0.5	$p = 0.995$ $x = 1.25$	$p = 0.999$ $x = 1.25$	$p = 0.9997$ $x = 1.25$	$p = 1$ $x = 1.25$	$p = 0.999$ $x = 0.86$	$p = 0.999$ $x = 0.779$	$p = 0.99$ $x = 0.89$ (ha $\lambda = 0.6$)	$p = 0.99$ $x = 0.83$ (ha $\lambda = 0.6$)
0.9	$p = 0.755$ $x = 1.006$	$p = 0.841$ $x = 1.006$	$p = 0.867$ $x = 1.006$	$p = 0.9636$ $x = 1.006$	$p = 0.9$ $x = 0.955$ $p = 0.999$ $x < 1$	$p = 0.95$ $x = 0.961$	$p = 0.95$ $x = 1.01$ ($\lambda = 0.905$-re)	$p = 0.95$ $x = 0.97$ (ha $\lambda = 0.905$)
			Reeves adatai					

Jól látható, hogy amíg a normális közelítésből illetve az (1) becslésből adódó felső határok még stacionárius esetben is ($\lambda = 0.5$, $\lambda = 0.9$) lehetnek 1-nél lényegesen nagyobbak, addig a folytonos közelítés eloszlásában a megfelelő felső határok mindig 1 alatt maradnak. Meg kell jegyezni, hogy a maximum likelihood becslés jósága a stacionárius esetben, és eloszlásának jó egyezése a folytonos eloszlásból számított közelítéssel szimulációs eredmények vizsgálata alapján, 20-40 megfigyelés mellett, már megalapozottnak tekinthető (ld. [3]!).

Álatalában az (1) becslés a stacionárius esetben az 1-hez közeli λ értékekre igen rossz. Ezt láthatjuk a [6] dolgozat 3. táblázatát tekintve, ahol az (1) becslésből 11 megfigyelés esetén:

ha $\lambda = 0.9$ akkor $p(\hat{\lambda} < 1) = 0.945$,

ha $\lambda = 0.95$ akkor $p(\hat{\lambda} < 1) = 0.773$,

míg a [2] táblázatából (a folytonos folyamat paraméterbecsléséből):

$$p_{\lambda=0.9}(\hat{\lambda} < 0.99) > 0.999 \qquad p_{\lambda=0.95}(\hat{\lambda} < 0.995) > 0.999.$$

Irodalom

[1] R.L. Anderson, (1942) Distribution of the serial correlation coefficient. Ann. Math. Statist., 13., 1-13.

[2] M. Арато – А. Бенцур: Функция распределения оценки параметра затухания стационарного гауссовского-марковского процесса, (1970) Studia Sci.Math. Hung., 5., 445-456.

[3] – (1972) Szimulációs eredmények az elemi Gauss folyamat paraméterei becsléseinek eloszlására, SzK Közlemények, 8., 3-35.

[4] J.P. Imhof, (1961) Computing the distribution of quadratic forms in normal variables, Biometrika, 48., 419-426.

[5] H.B. Mann, A. Wald, (1943) On the statistical treatment of linear stochastic difference equations. Econometrica, 11., 173-220.

[6] J.E. Reeves, (1972) The distribution of the maximum likelihood estimator of the parameter in the first-order autoregressive series, Biometrika, 59., 387-394.

[7] F.E. Satterthwaite, (1941) Synthesis of variance. Psychometrika, 6., 309-316.

[8] J.S. White (1958) The limiting distribution of the serial correlation coefficient in the explosive case, Ann. Math. Statist. 29., 1188-1197.

Beérkezett: 1973. szeptember 19.

Summary

In Reeves' paper [6] the exact distribution of estimator (1) is investigated. In the discrete time case the exact distribution may not be get for all values of λ, we may get only good approximation. The distribution is known for the maximum likelihood estimation of dumping parameter ρ (where $\lambda = e^{-\rho\Delta}$) of first order autoregressive gaussian stationary process with continuous time (see [2], [3]). The author shows on the basis of [3], that in the discrete time case estimate (1) may not be used when $\lambda \sim 1$ and the process is stationary. Further, from results [2] and [3] it follows that the approximation by continuous time to the discrete case is good with 20 obseratious. The author compares the results of Reeves with those which we may get from [2].

Р е з ю м е

В статье Reeves [6] рассматривается точное распределение оценки (1). В дискретном случае распределение не определяется при любом значении λ, но можно получить хорошое приближение. Известно распределение оценки наибольшего правдоподобия параметра затухания ρ (где $\lambda = e^{-\rho\Delta}$) стационарного гауссовского процесса авторегрессионного типа первого порядка с непрерывным временем (см. [2], [3]). Автор на основе [3], показывает, что в дискретном случае оценку (1) нельзя использовать при $\lambda \sim 1$ в стационарном случае. Далее, из результатов [2] и [3] следует, что приближение с непрерывным временем, для дискретного случая, хорошое уже при 20 наблюдений. Автор сравнивает результаты Reeves с результатами [2].

MTA Számítástechnikai és Automatizálási Kutató Intézete, **Közlemények** 11. (1973)

A KÉTLÉPCSŐS SZTOCHASZTIKUS PROGRAMOZÁSI PROBLÉMA EGY MEGOLDÁSI MÓDJÁRÓL

Strazicky Beáta

I.

A Dantzig és Madansky féle kétlépcsős sztochasztikus programozási probléma a következő:

$$\begin{aligned} (1) \qquad & \min z(\underline{x}) = \underline{c}'\underline{x} + E(\min \underline{q}'\underline{y} \mid T\underline{x} + M\underline{y} = \underline{\zeta}, \underline{y} \geqslant \underline{0}), \\ & A\underline{x} = \underline{b}, \\ & \underline{x} \geqslant \underline{0}, \end{aligned}$$

ahol A mxn, T m_0xn, M $m_0xn_0$méretű mátrixok, $\underline{\zeta} \in R_{m_0}$ véletlen vektor, $\underline{b} \in R_m$, $\underline{c} \in R_n$, $\underline{q} \in R_{n_0}$, $\underline{x} \in R_n$, $\underline{y} \in R_{n_0}$ vektorok és E várható értéket jelöl.

A

$$\begin{aligned} & \min \underline{q}'\underline{y} \\ & M\underline{y} = \underline{\zeta} - T\underline{x} \\ & \underline{y} \geqslant \underline{0} \end{aligned}$$

problémát, ahol $\underline{x}$ rögzített, második lépcső problémának nevezzük.
Megengedett egy $\underline{x}$ vektor az (1) feladatban, ha $A\underline{x} = \underline{b}$, $\underline{x} \geqslant \underline{0}$ és minden lehetséges $\underline{\zeta}$ esetén a második lépcső problémának van optimális megoldása.

Tegyük fel, hogy a $\underline{\zeta}$ valószínűségi vektorváltozó véges sok lehetséges értéket vehet fel, legyenek ezek $\underline{f}_1, \underline{f}_2, \dots, \underline{f}_k$, a megfelelő valószínűségek pedig $p_1, p_2, \dots, p_k$.
Ekkor a kétlépcsős probléma a következő feladattal ekvivalens:

$$\begin{aligned} & \min(\underline{c}'\underline{x} + p_1\underline{q}'\underline{y}_1 + \dots + p_k\underline{q}'\underline{y}_k), \\ (2) \qquad & A\underline{x} \phantom{+ M\underline{y}_1 + M\underline{y}_2 + \dots + M\underline{y}_k} = \underline{b}, \\ & T\underline{x} + M\underline{y}_1 = \underline{f}_1, \\ & T\underline{x} + \phantom{M\underline{y}_1 +} M\underline{y}_2 = \underline{f}_2, \\ & \quad \vdots \qquad\qquad \ddots \qquad\qquad \vdots \\ & T\underline{x} + \phantom{M\underline{y}_1 + M\underline{y}_2 + \dots +} M\underline{y}_k = \underline{f}_k, \\ & \underline{x} \geqslant \underline{0}, \underline{y}_1 \geqslant \underline{0}, \dots, \underline{y}_k \geqslant \underline{0}. \end{aligned}$$

E lineáris programozási feladat duálisának megoldására fogunk egy módszert kidolgozni, amely a megoldást nagy k érték esetén is lehetővé teszi.

A (2) feladat duálisa a következő:

$$
\begin{aligned}
&\max(\underline{b}'\underline{u} + \underline{f}_1'\underline{v}_1 + \ldots + \underline{f}_k'\underline{v}_k)\,, \\
(3)\qquad &A'\underline{u} + T'\underline{v}_1 + \ldots + T'\underline{v}_k \leqslant \underline{c}\,, \\
&M'\underline{v}_1 \qquad\qquad \leqslant p_1\underline{q}\,, \\
&\qquad\vdots \\
&\qquad M'\underline{v}_k \leqslant p_k\underline{q}\,,
\end{aligned}
$$

Vezessük be a $\underline{w}_i = \frac{1}{p_1}\underline{v}_i$, $i = 1, 2, \ldots, k$ új változókat és legyen $\underline{u} = \underline{u}^+ - \underline{u}^-$, $\underline{w}_i = \underline{w}_i^+ - \underline{w}_i^-$, $i = 1, 2, \ldots, k$, ahol $\underline{u}^+, \underline{u}^- \geqslant \underline{0}$, és $\underline{w}_i^+, \underline{w}_i^- \geqslant \underline{0}$, $i = 1, 2, \ldots, k$. Az $\underline{u}^+, \underline{u}^-, \underline{w}_i^+, \underline{w}_i^-$, $i = 1, 2, \ldots, k$ változókban megfogalmazott duálfeladat így a következő lesz:

$$
\begin{aligned}
&\max(p_1\underline{f}_1'\underline{w}_1^+ - p_1\underline{f}_1'\underline{w}_1^- + \ldots + p_k\underline{f}_k'\underline{w}_k^+ - p_k\underline{f}_k'\underline{w}_k^- + \underline{b}'\underline{u}^+ - \underline{b}'\underline{u}^-)\,, \\
(4)\qquad &p_1T'\underline{w}_1^+ - p_1T'\underline{w}_1^- + \ldots + p_kT'\underline{w}_k^+ - p_kT'\underline{w}_k^- + A'\underline{u}^+ - A'\underline{u}^- \leqslant \underline{c}\,, \\
&M'\underline{w}_1^+ - M'\underline{w}_1^- \qquad \leqslant \underline{q}\,, \\
&\qquad\vdots \\
&\qquad M'\underline{w}_k^+ - M'\underline{w}_k^- \leqslant \underline{q}\,, \\
&\qquad \underline{w}_i^+, \underline{w}_i^- \geqslant \underline{0}\,, \quad i = 1, 2, \ldots, k\,, \ \underline{u}^+ \geqslant \underline{0}\,, \ \underline{u}^- \geqslant \underline{0}\,.
\end{aligned}
$$

A (4) feladat feltételeit slack változók bevezetésével egyenlőséggé alakítva a kapott feladat szimplex módszerrel megoldható. A megoldás azonban nagy k esetén a rendelkezésre álló szimplex algoritmusokkal nem lehetséges nagy memóriaigény és számítási idő igény miatt. Ha Bakes [2] gondolatmenetét alkalmazzuk e speciális feladatra, egy olyan bázisdekompozíciós algoritmust kapunk, amelynek felhasználása esetén a fenti problémák nem lépnek fel.

II.

Alkalmazzuk Bakes gondolatmenetét a

$$
\begin{aligned}
&\max(p_1\underline{f}_1'\underline{w}_1^+ - p_1\underline{f}_1'\underline{w}_1^- + \ldots + p_k\underline{f}_k'\underline{w}_k^+ - p_k\underline{f}_k'\underline{w}_k^- + \underline{b}'\underline{u}^+ - \underline{b}'\underline{u}^-)\,, \\
&M'\underline{w}_1^+ - M'\underline{w}_1^- + I\underline{s}_1 \qquad = \underline{q}, \\
&\qquad\vdots \\
&\qquad M'\underline{w}_k^+ - M'\underline{w}_k^- + I\underline{s}_k = \underline{q}, \\
&p_1T'\underline{w}_1^+ - p_1T'\underline{w}_1^- + \ldots + p_kT'\underline{w}_k^+ - p_kT'\underline{w}_k^- + A'\underline{u}^+ - A'\underline{u}^- + I\underline{s} = \underline{c}, \\
&\underline{w}_i^+, \underline{w}_i^-, \underline{s}_i \geqslant \underline{0}, \quad i = 1, 2, \ldots, k, \quad \underline{u}^+, \underline{u}^-, \underline{s} \geqslant \underline{0},
\end{aligned}
$$

lineáris programozási feladat megoldására, ahol I megfelelő méretű egységmátrix. Tekintsük ennek során alapfeladatnak a

$$\max(p_1\underline{f}_1'\underline{w}_1^+ - p_1\underline{f}_1'\underline{w}_1^- + \ldots + p_k\underline{f}_k'\underline{w}_k^+ - p_k\underline{f}_k'\underline{w}_k^-),$$

$$(6)\qquad \begin{aligned} M'\underline{w}_1^+ - M'\underline{w}_1^- + I\underline{s}_1 &= \underline{q}, \\ &\;\;\vdots \\ M'\underline{w}_k^+ - M'\underline{w}_k^- + I\underline{s}_k &= \underline{q}, \end{aligned}$$

$$\underline{w}_i^+,\ \underline{w}_i^-,\ \underline{s}_i \geqslant \underline{0},\quad i = 1, 2, \ldots, k,$$

feladatot, addicionális feltételeknek pedig a

$$(7)\qquad p_1T'\underline{w}_1^+ - p_1T'\underline{w}_1^- + \ldots + p_kT'\underline{w}_k^+ - p_kT'\underline{w}_k^- + A'\underline{u}^+ - A'\underline{u}^- + I\underline{s} = \underline{c},$$

$$\underline{u}^+, \underline{u}^-, \underline{s} \geqslant \underline{0}$$

feltételeket.

A gondolatmenet a következő:

Megmutatjuk, hogy

a/ az (5) feladat egy megengedett bázisa speciális felépítésű,

b/ ilyen speciális felépítésű megengedett bázis esetén a megfelelő szimplex iterációhoz szükséges információk a teljes bázis-inverz ismerete nélkül, egyszerűen számíthatók,

c/ a báziscsere végezhető úgy is, hogy a bázis speciális felépítése ne változzon.

Ha a szimplex iterációk sorát egy ilyen speciális felépítésű bázis vizsgálatával kezdjük, majd minden iterációban a bázis felépítését változatlanul hagyjuk, akkor minden iteráció végezhető a fenti "egyszerű" módon. Igy a szimplex módszer egy, a problémára specializált változatát nyerjük. Megjegyzés: ugyanez a gondolata a [2], [3], [4], [5], [6] cikkekben szereplő eljárásoknak.

a/ Tegyük fel, hogy a (6) feladatnak van megengedett megoldása. Ha a (6)-nak nincs megengedett megoldása, akkor (5)-nek sincs. Legyen B (6) egy megengedett bázisa. Feltehető, hogy ez

$$B = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{pmatrix}$$

felépítésű, ahol a $B_1, \ldots, B_k$ blokkokon kívül csupa nulla szerepel, B_i pedig az

$$(8)\qquad M'\underline{w}^+ - M'\underline{w}^- + I\underline{s} = \underline{q},\quad \underline{w}^+ \geqslant \underline{0},\quad \underline{w}^- \geqslant \underline{0},\quad \underline{s} \geqslant \underline{0}$$

megengedett bázisa, minden $i = 1, 2, \ldots, k$ esetén.

Nevezzük a továbbiakban az (5) illetve (6) feladat együttható mátrixának i-edik oszlopvektorát

$$(l-1)(2m_0 + n_0) < i \leq l(2m_0 + n_0)$$

teljesülése esetén az l-edik lépcsőfokhoz tartozónak, ahol $1 \leq l \leq k$, illetve az (5) feladat i-edik oszlopvektorát a 0-adik lépcsőfokhoz tartozónak, ha

$$K(2m_0 + n_0) < i \leq K(2m_0 + n_0) + 2m + n .$$

Előző feltevésünk megfogalmazható úgy is, hogy (6) egy megengedett bázisa olyan, hogy az i-edik lépcsőfokhoz tartozó vektorok közül tartalmaz olyanokat és csak olyanokat, melyeknek j-edik komponenseiből, ahol $(i-1)n_0 < j \leq in_0$, alkotott vektorok (8) megengedett bázisát alkotják és a vektorok sorrendje olyan, hogy ez a bázis szerepel B i-edik blokkjában.
Legyen

$$\hat{B} = \begin{pmatrix} B & 0 \\ L & I^* \end{pmatrix}$$

ahol az L mátrixot a (7) együtthatómátrixának a B-beli vektorokkal azonos indexű oszlopvektorai alkotják, az I^* mátrix oszlopvektorait a következő módon választjuk:

Legyen $\underline{x}_{B1}$ a $B\underline{x}_{B1} = \begin{pmatrix} \underline{q} \\ \vdots \\ \underline{q} \end{pmatrix}$ egyenlőségrendszer megoldásvektora, $\underline{x}_{B1} \geq \underline{0}$ mivel B megengedett bázisa (6)-nak;
és legyen $\underline{t} = L\underline{x}_{B1} - \underline{c}$. Jelölje t_i a $\underline{t}$ vektor i-edik komponensét.
Ha $t_i \geq 0$, akkor válasszuk I^* i-edik oszlopául az i-edik n-dimenziós egységvektort, ha pedig $t_i < 0$, akkor az i-edik n-dimenziós egységvektor -1 szeresét, $i = 1, 2, \ldots, n$. Ha minden $t_i \geq 0$, $i = 1, 2, \ldots, n$, akkor $\hat{B}$ (5) egy megengedett bázisa. Ha nem, akkor $\hat{B}$ az 5-ből, a $t_i < 0$ komponensekhez tartozó

$$\begin{pmatrix} \underline{0} \\ -\underline{e}_i \end{pmatrix},$$

$\underline{0} \epsilon R_{kn_0}$ vektorokkal bővített feladat egy megengedett bázisa. Ha van ilyen komponens, akkor eljárásunk során $\hat{B}$-ból kiindulva egy első fázisnak megfelelő célfüggvényt kell először maximalizálnunk.

Tegyük fel, hogy a szimplex iteráció során vizsgálandó megengedett bázis

(9)
$$\hat{B} = \begin{pmatrix} B & Y \\ L & Z \end{pmatrix}$$

felépítésű, ahol B (6) egy bázisa, ezért feltehetően

(9a) $$B = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{pmatrix}$$

felépítésű. Feltevésünk mindenesetre az első iterációban az előzőek miatt teljesül, mégpedig $Y = 0$ és $Z = I^*$. A továbbiakban:

Y a (6) együtthatómátrix $\hat{B}$-ban szereplő további oszlopvektoraiból

Z a (7), az Y-ban szereplő oszlopvektorokkal azonos indexű oszlopvektoraiból áll.

Feltesszük még, hogy Y-ban az oszlopvektorok a lépcsőfoknak megfelelő sorrendben követik egymást, eltekintve a 0. lépcsőfokhoz tartozó bázisbeli vektoroktól, melyek Y utolsó oszlopai. Mivel B bázis, $Y = BX$ alakban írható, s az oszlopok sorrendje miatt

(9b) $$X = \begin{pmatrix} X_1 & & & \\ & \ddots & & 0 \\ & & X_k & \end{pmatrix}$$

ahol az X mátrix valamely X_i blokkja esetleg egyetlen oszlopot sem tartalmaz. (Akkor ugyanis, ha a megfelelő lépcsőfokhoz tartozó vektorok közül csak n_0 szerepel a bázisban.)

Bontsuk B és Y particiójának megfelelően L-et és Z-t is blokkokra.
Legyen

$$L = (p_1 L_1, \ldots, p_k L_k),$$
$$Z = (p_1 Z_1, \ldots, p_k Z_k, Z_0)$$

ahol L_i, $(i = 1, 2, \ldots, k)$ a $(T', -T', 0)$ mátrix B_i bázisbeli vektorokkal azonos indexű oszlopvektorait tartalmazza, Z_i, $(i = 1, 2, \ldots, k)$ pedig e mátrixnak X_i-beli vektorokkal azonos indexű vektorait tartalmazza.

Z_0 oszlopai a bázisbeli, 0-adik lépcsőfokhoz tartozó vektorokkal azonos indexű oszlopok az $(A', -A', I)$ mátrixból.

Feltevésünk szerint tehát:

(10) $$\hat{B} = \left(\begin{array}{ccc|ccc} B_1 & & & B_1 X_1 & & \\ & \ddots & & & \ddots & \\ & & B_k & & & B_k X_k \\ \hline & p_1 L_1, \ldots, p_k L_k & & & p_1 Z_1, \ldots, p_k Z_k, Z_0 & \end{array}\right)$$

felépítésű.

b/ Jelölje $\underline{\rho}$ az (5) feladat célfüggvényegyütthatóinak bázisbeli komponenseiből álló vektort. Bontsuk ezt $\underline{\rho}^T = (\underline{\rho}_1^T, \underline{\rho}_2^T)$ részekre a $\hat{B}$ (9) felbontásának megfelelően. $\hat{B}$ optimális bázis, ha a

(11) $\quad (\underline{\pi}_1', \underline{\pi}_2')\hat{B} = \underline{\rho}'$

egyenlőségrendszer $(\underline{\pi}_1', \underline{\pi}_2')$ megoldására, (ahol: $\underline{\pi}_1 \epsilon R_{k \cdot n_0}$, $\underline{\pi}_2 \epsilon R_n$),

(12)
$$(\underline{\pi}_1', \underline{\pi}_2') \begin{pmatrix} M', & -M', & I, & & & & \\ & & & \ddots & & & \\ & & & & M', & -M', I, & \\ p_1 T', & -p_1 T', 0, & \ldots, & p_k T', & -p_k T', 0, A', & -A', I \end{pmatrix} \geqslant$$

$$\geqslant (p_1 \underline{f}_1', -p_1 \underline{f}_1', \underline{0}', \ldots, p_k f_k', -p_k f_k', \underline{0}', \underline{b}', -\underline{b}', \underline{0}')$$

teljesül.

Határozzuk meg a $(\underline{\pi}_1', \underline{\pi}_2')$ vektort (9) felbontás felhasználásával a következők alapján:

(11) a következő alakú:

(13)
$$\underline{\pi}_1' B + \underline{\pi}_2' L = \underline{\rho}_1'$$
$$\underline{\pi}_1' BX + \underline{\pi}_2' Z = \underline{\rho}_2'$$

innen

(14)
$$\underline{\pi}_1' BX + \underline{\pi}_2' LX = \underline{\rho}_1' X$$
$$\underline{\pi}_1' BX + \underline{\pi}_2' Z = \underline{\rho}_2'$$

ebből pedig

$$\underline{\pi}_2'(LX - Z) = \underline{\rho}_1' X - \underline{\rho}_2' .$$

Az $(LX - Z)$ mátrix oszlopainak száma megegyezik sorainak számával és oszlopai lineárisan függetlenek. Tegyük fel ugyanis, hogy nem lineárisan függetlenek, azaz létezik olyan $\underline{\lambda}$, $\underline{\lambda} \neq \underline{0}$, melyre:

$$(LX - Z)\underline{\lambda} = \underline{0} .$$

Legyen $\underline{\mu} = X\underline{\lambda}$. Ekkor $(\underline{\mu}', -\underline{\lambda}')$ vektorra:

$$\begin{pmatrix} B & BX \\ L & Z \end{pmatrix} \begin{pmatrix} \underline{\mu} \\ -\underline{\lambda} \end{pmatrix} = \underline{0}$$

ami ellentmondás azzal, hogy $\hat{B}$ bázis. $(LX - Z)$ tehát invertálható, és így:

$$(15) \qquad \underline{\pi}_2' = (\underline{\rho}_1' X - \underline{\rho}_2')(LX - Z)^{-1} ,$$

$$\underline{\pi}_1' = (\underline{\rho}_1' - \underline{\pi}_2' L)B^{-1} .$$

Ezért: $(\underline{\pi}_1', \underline{\pi}_2')$ meghatározásához k $(n_0 \times n_0)$ méretű mátrix inverzének és egy $(n \times n)$ méretű mátrix inverzének ismerete szükséges.

Legyen $\underline{\pi}_1' = (\underline{\pi}_{11}', \ldots, \underline{\pi}_{1k}')$, $\underline{\pi}_{1i} \epsilon R_{n_0}$, $i = 1, 2, \ldots, k$. Ha a $(\underline{\pi}_1', \underline{\pi}_2')$ vektorra teljesülnek a

$$(16) \qquad \underline{\pi}_{1i}' M' + p_i \underline{\pi}_2' T' = p_i \underline{f}_i', \qquad i = 1, 2, \ldots, k .$$

$$(17) \qquad \underline{\pi}_{1i} \geqslant \underline{0}, \qquad i = 1, 2, \ldots, k ,$$

$$(18) \qquad \underline{\pi}_2' A' = \underline{b}' ,$$

$$(19) \qquad \underline{\pi}_2 \geqslant \underline{0}$$

feltételek, akkor a $\hat{B}$ bázis optimális.

Egyébként:

1. ha van olyan i és j index, $1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant m$, melyre

$$(\underline{\pi}_{1i}' M' + p_i \underline{\pi}_2' T')_j < p_i \underline{f}_{ij}$$

vagy

$$(\pi_{1i}' M' + p_i \underline{\pi}_2' T')_j > p_i \underline{f}_{ij}$$

azaz

$$\underline{\widetilde{\pi}}_{1l} = \frac{\underline{\pi}_{1l}}{p_l}, \qquad l = 1, 2, \ldots, k$$

jelölésekkel:

$$(\underline{\pi}_2' T')_j < (\underline{f}_i^T - \underline{\widetilde{\pi}}_{1i}' M')_j$$

vagy

$$(\underline{\pi}_2' T')_j > (\underline{f}_i' - \underline{\widetilde{\pi}}_{1i} M')_j$$

akkor a bázisba vonandó vektor:

$$\begin{pmatrix} \underline{0}_1 \\ \underline{m}_j \\ \underline{0}_2 \\ p_i \underline{t}_j \end{pmatrix} \quad \text{ill.} \quad \begin{pmatrix} \underline{0}_1 \\ -\underline{m}_j \\ \underline{0}_2 \\ -p_i \underline{t}_j \end{pmatrix}$$

ahol $\underline{0}_1 \in R_{(i-1)n_0}$, $\underline{0}_2 \in R_{(k-1)n_0}$ zéró vektor, $\underline{m}_j$ ill. $\underline{t}_j$ az M' ill. T' mátrix j-edik oszlopa.

2. ha van olyan i és j index, $1 \leq i \leq k$, $1 \leq j \leq n_0$ melyre: $(\underline{\pi}_{1i})_j < 0$ akkor a bevonandó vektor:

$$\begin{pmatrix} \underline{0}_1 \\ \underline{e}_j \\ \underline{0}_2 \end{pmatrix}$$

, ahol $\underline{0}_1 \in R_{(i-1)n_0}$ és $\underline{0}_2 \in R_{(k-1)n_0+n}$ zéró vektor, $\underline{e}_j \in R_{n_0}$, j-edik egységvektor.

3. ha van olyan j index, $1 \leq j \leq m$, melyre $(\underline{\pi}_2' A')_j < (\underline{b})_j$ vagy $(\underline{\pi}_2' A')_j > (\underline{b})_j$ akkor a bevonandó vektor

$$\begin{pmatrix} \underline{0} \\ \underline{a}_j \end{pmatrix} \quad \text{ill.} \quad \begin{pmatrix} \underline{0} \\ -\underline{a}_j \end{pmatrix}$$

ahol $\underline{0} \in R_{k \cdot n_0}$, $\underline{a}_j$ pedig A' j-edik oszlopa.

4. ha van olyan j index, $1 \leq j \leq n$ melyre: $(\underline{\pi}_2)_j < 0$, akkor a bevonandó vektor

$$\begin{pmatrix} \underline{0} \\ \underline{e}_j \end{pmatrix}$$

, ahol $\underline{0} \in R_{k \cdot n_0}$, $\underline{e}_j \in R_n$: j-edik egységvektor.

Ha a bázis nem optimális, akkor a bevonandó vektor meghatározását ennek az aktuális bázisban történő előállítása, illetve az aktuális bázishoz tartozó bázismegoldás meghatározása követi.

Használjuk ennek során a következő jelöléseket:

Legyen $\underline{d}$ a bevonandó vektor előállítása az aktuális bázisban, ill. $\underline{x}_{\hat{B}}$ a megfelelő bázismegoldás.

Bontsuk fel $\underline{d}$-t és $\underline{x}_{\hat{B}}$-t a $\hat{B}$ bázis (9) felbontásának megfelelően. Legyen

$$\underline{d}' = (\underline{d}_1', \underline{d}_2') \quad \text{ill.} \quad \underline{x}_{\hat{B}}' = (\underline{x}_{\hat{B}1}', \underline{x}_{\hat{B}2}')$$

ahol tehát: $\underline{d}_1$ és $\underline{x}_{\hat{B}1} \in R_{k \cdot n_0}$ és $\underline{d}_2$, $\underline{x}_{\hat{B}2} \in R_n$.

Bontsuk tovább $\underline{d}_1$-et és $\underline{x}_{\hat{B}1}$-et

$$\underline{d}_1' = (\underline{d}_{11}', \ldots, \underline{d}_{1k}') \quad \text{ill.} \quad \underline{x}_{\hat{B}1}' = (\underline{x}_{\hat{B}1,1}', \ldots, \underline{x}_{\hat{B}1,k}')$$

részekre, ahol: $\underline{d}_{1i}$, $\underline{x}_{\hat{B}1,i} \in R_{n_0}$, $i = 1, 2, \ldots, k$, és legyen $\underline{d}_2$ ill. $\underline{x}_{\hat{B}2}$ (9b)-vel azonos módon felbontva:

$$\underline{d}_2' = (\underline{d}_{21}', \ldots, \underline{d}_{2k}', \underline{d}_{20}'), \qquad \underline{x}_{\hat{B}2}' = (\underline{x}_{\hat{B}2,1}', \ldots, \underline{x}_{\hat{B}2,k}', \underline{x}_{\hat{B}2,0}')$$

Meghatározandó tehát először az a $(\underline{d}_1', \underline{d}_2')$ melyre:

(20) $B\underline{d}_1 + BX\underline{d}_2 = \underline{u}$

(21) $L\underline{d}_1 + Z\underline{d}_2 = \underline{v}$

ahol $\begin{pmatrix} \underline{u} \\ \underline{v} \end{pmatrix}$ a bevonandó vektor.

(20)-ból: $B(\underline{d}_1 + X\underline{d}_2) = \underline{u}$.

Legyen $\underline{d}_0 = \underline{d}_1 + X\underline{d}_2$, ekkor: $\underline{d}_0 = B^{-1}\underline{u}$ és $\underline{d}_1 = \underline{d}_0 - X\underline{d}_2 - t$ (21)-be behelyettesítve:

$$L\underline{d}_0 - LX\underline{d}_2 + Z\underline{d}_2 = \underline{v}$$

azaz:

$$\underline{d}_2 = (LX - Z)^{-1}(L\underline{d}_0 - \underline{v}) .$$

Most vizsgáljuk $\underline{d}_0, \underline{d}_1, \underline{d}_2$ értékét a bevonandó vektor négy lehetséges típusa szerint:

1/a $\underline{u}' = (\underline{0}', \ldots, \underline{0}', \underline{m}_j', \underline{0}', \ldots, \underline{0}')$, $\underline{v} = p_i \underline{t}_j$

$$\underline{d}_0 = \begin{pmatrix} B_1^{-1} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & & & & B_k^{-1} \end{pmatrix} \begin{pmatrix} \underline{0}^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ \underline{0}^{(i-1)} \\ \underline{m}_j \\ \underline{0}^{(i+1)} \\ \underline{0}^{(k)} \end{pmatrix}$$

Legyen:

$$\underline{d}_0' = (\underline{d}_{01}', \ldots, \underline{d}_{0k}') , \quad \underline{d}_{0l} \in R_{n_0} , \quad l = 1, 2, \ldots, k .$$

$$\underline{d}_{0t} = \underline{0} \qquad t = 1, 2, \ldots, k , \quad t \neq i .$$

$$\underline{d}_{0i} = B_i^{-1} \underline{m}_j$$

$$L\underline{d}_0 = p_i L_i B_i^{-1} \underline{m}_j$$

$$L\underline{d}_0 - \underline{v} = p_i (L_i B_i^{-1} \underline{m}_j - \underline{t}_j)$$

$$\underline{d}_2 = p_i (LX - Z)^{-1} (L_i B_i^{-1} \underline{m}_j - \underline{t}_j) ,$$

$$\underline{d}_1 = \underline{d}_0 - X\underline{d}_2 .$$

1/b $\underline{u} = (\underline{0}, \ldots, \underline{0}, -\underline{m}_j, \underline{0}, \ldots, \underline{0})$, $\underline{v} = -p_i \underline{t}_j$

alapján:

$$d_{0t} = 0 \quad t = 1, 2, \ldots, k\,, \quad t \neq i\,, \tag{22}$$

$$\underline{d}_{0i} = -B_i^{-1} \underline{m}_j\,,$$

$$\underline{d}_2 = -p_i (LX - Z)^{-1} (L_i B_i^{-1} \underline{m}_j - \underline{t}_j)\,,$$

$$\underline{d}_1 = \underline{d}_0 - X\underline{d}_2\,.$$

2. esetben:

$$\underline{u}' = (\underline{0}', \ldots, \underline{0}', \underline{e}_j, \underline{0}', \ldots, \underline{0}')\,, \quad \underline{v} = \underline{0} \tag{23}$$

$$\underline{d}_{0t} = \underline{0}\,, \quad t = 1, 2, \ldots, k\,, \quad t \neq i\,,$$

$$\underline{d}_{0i} = B_i^{-1} \underline{e}_j\,,$$

$$\underline{d}_2 = p_i (LX - Z)^{-1} L_i B_i^{-1} \underline{e}_j\,,$$

$$\underline{d}_1 = \underline{d}_0 - X\underline{d}_2\,, \quad \underline{d}_{1t} = \underline{d}_{0t} - X_t \underline{d}_{2t}\,, \quad t = 1, 2, \ldots, k\,.$$

3. esetben: a) $\underline{u} = \underline{0}$, $\underline{v} = \underline{a}_j$

$$\underline{d}_0 = \underline{0} \tag{24}$$

$$\underline{d}_2 = -(LX - Z)^{-1} \underline{a}_j$$

$$\underline{d}_1 = X(LX - Z)^{-1} \underline{a}_j$$

$$\underline{d}_{1t} = X_t \underline{d}_{2t}\,, \quad t = 1, 2, \ldots, k\,.$$

b) $\underline{u} = \underline{0}$, $\underline{v} = -\underline{a}_j$, így $\underline{d}_1, \underline{d}_2$ az előző -1 szerese.

4. esetben:

$\underline{u} = \underline{0}$, $\underline{v} = \underline{e}_j$

$$\underline{d}_0 = \underline{0}\,, \quad \underline{d}_2 = -(LX - Z)^{-1} \underline{e}_j\,, \quad \underline{d}_1 = X(LX - Z)^{-1} \underline{e}_j\,. \tag{25}$$

Az $(\underline{x}'_{\hat{B}1}, \underline{x}'_{\hat{B}2})$ megoldásvektor esetén: $\underline{u}' = (\underline{q}', \ldots, \underline{q}')$, $\underline{v} = \underline{c}$.

Legyen $\underline{x}_0 = \underline{x}_{\hat{B}1} + X\underline{x}_{\hat{B}2}$, $\underline{x}_0 \in R_{kn_0}$. Ekkor:

$$\underline{x}'_0 = (\underline{x}'_{01}, \ldots, \underline{x}'_{0k})\,, \quad \underline{x}_{0i} \in R_{n_0}\,, \quad i = 1, 2, \ldots, k\,. \tag{26}$$

$$\underline{x}_{0l} = B_l^{-1} \underline{q}\,, \quad l = 1, 2, \ldots, k\,.$$

$$\underline{x}_{\hat{B}2} = (LX - Z)^{-1}(L\underline{x}_0 - \underline{c}),$$

$$\underline{x}_{\hat{B}1} = \underline{x}_0 - X\underline{x}_{\hat{B}2},$$

$$L\underline{x}_0 - \underline{c} = \sum_{i=1}^{k} p_i L_i B_i^{-1}\underline{q} - \underline{c},$$

$$\underline{x}_{\hat{B}1,i} = B_i^{-1}\underline{q} - X_i\underline{x}_{\hat{B}2,i}, \quad i = 1, 2, \ldots, k.$$

c/ Megmutatjuk, hogy a báziscsere végezhető úgy is, hogy a bázis felépítése ne változzék, s hogy ekkor egy-egy iteráció során a szükséges inverz-mátrixok közül legfeljebb kettő változik meg, valamely B_i^{-1} és az $(LX - Z)^{-1}$.

Vizsgáljuk a következő lehetséges eseteket:

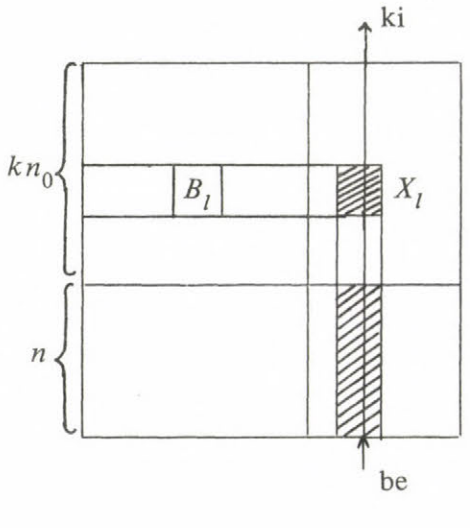
I. A bázist elhagyó vektor az együtthatómátrix l-edik lépcsőfokához tartozik, $1 \leq l \leq k$.

I/1 a bázist elhagyó vektor a bázis utolsó n oszlopvektora között szerepel.

I/2 a bázist elhagyó vektor a bázis első $k \cdot n_0$ oszlopvektora között szerepel.

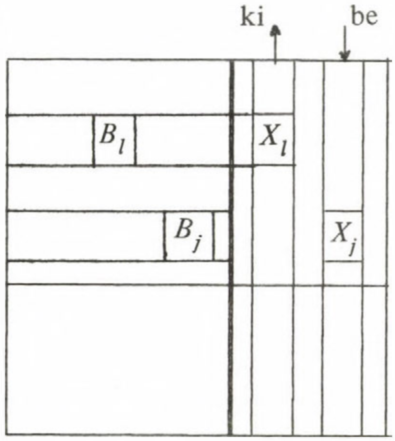
II. A bázist elhagyó vektor az együtthatómátrix 0. lépcsőfokához tartozik.

I/1 esetben:

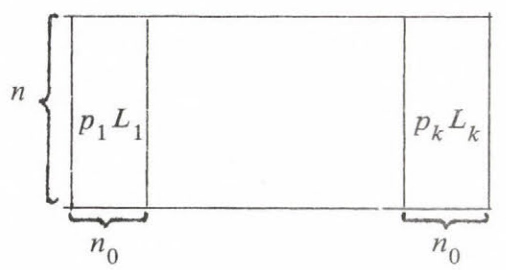


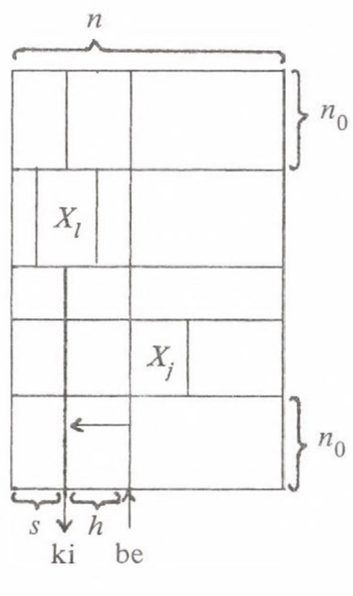
a) ha a bevonandó vektor is az l-edik lépcsőfokhoz tartozik, akkor a két vektor kicserélése a bázis felépítését nem változtatja meg. Ha a kilépő vektor az X_l blokkban az s-edik helyen szerepel, akkor a báziscserénél csak az X_l és Z_l blokkok s-edik oszlopa és ezek miatt az $(LX - Z)$ mátrix s-edik oszlopa változik meg. Mégpedig:

Ha a bevonandó vektor l-edik lépcsőfokhoz tartozó része $\underline{m}$, (7) együtthatómátrixhoz tartozó része pedig $\underline{t}$, akkor $(LX - Z)$ s-edik oszlopvektora $\underline{r} = L_l B_l^{-1}\underline{m} - \underline{t}$ lesz.



b) ha a bevonandó vektor a j-edik lépcsőfokhoz tartozik, $1 \leqslant j \leqslant k$, $j \neq l$, akkor végezzük a báziscserét úgy, hogy az X_l blokkhoz tartozó vektorok közül hagyjuk el a bázisból kilépőt, a bevonandó vekvektort pedig X_j blokkba vonjuk. Ekkor tehát az X_l blokk oszlopainak száma csökken, az X_j blokk oszlopainak száma növekszik, a bázis felépítése pedig változatlan. $(LX - Z)$ a következő módon változik:





ha a bázisból kilépő vektor az X blokk s-edik vektora, a bázisba vonandó vektor pedig az X blokk $s + h$-adik helyére kerül, akkor, ha $h > 0$ akkor:

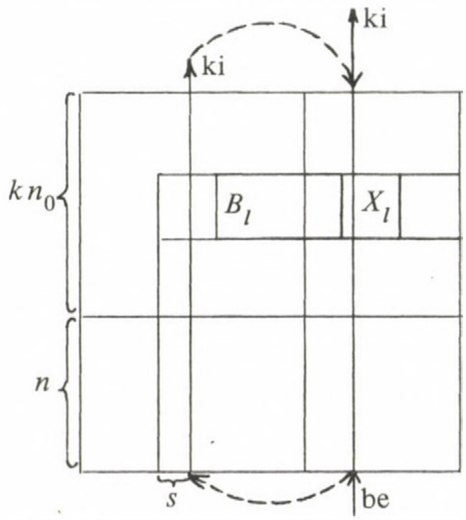
az $(LX - Z)$ első $s - 1$ oszlopa változatlan, $s + i$-edik oszlopa helyére az $(s + i + 1)$-edik oszlopvektor kerül, $i = 1, 2, \ldots, h - 1$.

$(s + h)$-adik oszlopa $\underline{r} = L_j B_j^{-1} \underline{m} - \underline{t}$ lesz, ahol $\underline{m}$ a bevonandó vektor j-edik lépcsőfokához tartozó, $\underline{t}$ pedig a (7) együtthatómátrixhoz tartozó része. $(LX - Z)$ további oszlopai nem változnak.

$h < 0$ eset hasonlóan meggondolható.

c) A bevonandó vektor a 0. lépcsőfokhoz tartozik: ha a bevonandó vektort a Z_0 blokkba vonjuk az X_l blokk elemeinek száma csökken, a Z_0 blokk elemeinek száma nő, de a bázis felépítése nem változik. Ha a bázisból kilépő vektor az X-ben az s-edik oszlopvektor, a bevonandó vektor pedig az X $s+h$-adik helyére kerül, akkor $(LX - Z)$ ugyanúgy változik, mint az előző esetben, eltekintve attól, hogy $s+h$-adik oszlopa: $-\underline{t}$ lesz.

1/2 eset: A bázist elhagyó vektor az l-edik lépcsőfokhoz tartozik és a bázisban az első $k \cdot n_0$ vektor között szerepel.



Tegyük fel, hogy a bázist elhagyó vektor a B_l blokk s-edik oszlopa. Két eset lehetséges:

a, X_l s-edik sorában van nem zéró érték

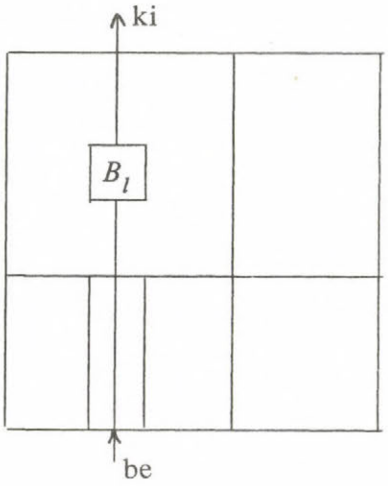
b, X_l s-edik sorában nincs nem zéró érték.

Végezzük a báziscserét az első esetben a következő módon:

Először cseréljük ki B_l s-edik oszlopvektorát X_l egy olyan vektorával, melynek s-edik komponense nem zéró.
E csere után kapott új B_l blokk ismét bázisa lesz (8)-nak, így e csere elvégzése a bázis felépítését nem módosítja.

A tulajdonképpeni báziscserét ezután már I/1 alapján végezhetjük, a bázis felépítése nem változik. A csere során a B_l, X_l, L_l, Z_l blokkok változnak a két vektor bázison belüli felcserélése miatt, majd a tulajdonképpeni báziscsere során I/1-nek megfelelően változnak a blokkok.

b. Ha X_l s-edik sorában nincs nem zéró érték, akkor a bázisba bevonandó vektor csakis az l-edik lépcsőfokhoz tartozó lehet, mégpedig olyan, hogy az l-edik lépcsőfokhoz tartozó része B_l további vektoraival (8) bázisát alkotja. [Megjegyzés (1) miatt.] Igy a két vektor felcserélése a bázis felépítését nem módosítja.



A csere során megváltozik a B_l és L_l blokk, B_l megváltozása miatt az X_l blokk, X_l és L_l megváltozása miatt az $(LX - Z)$ mátrix változik. Mégpedig:

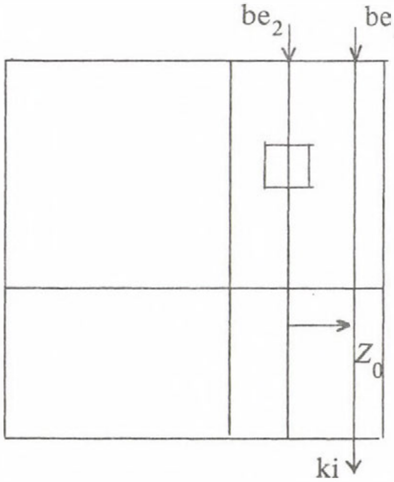
Ha E a B_l báziscsere során az inverz szorzat előállításához szükséges majdnem-egységmátrixot, F pedig az L_l s-edik oszlopának felcseréléséhez szükséges majdnem-zéró mátrixot jelöli (melynek tehát csak s-edik oszlopa nem zéró, és ennek értéke: $\underline{r} = \underline{t} - \underline{t}_0$, ahol:

$\underline{t}$ a bevonandó vektor 0-adik lépcsőfokhoz tartozó része

$\underline{t}_0$ pedig L_l s-edik oszlopa, akkor: az LX szorzatmátrix X_l-vel azonos indexű oszlopai változnak csak, az általuk alkotott blokk új értéke:

$$(L_l + F)EX_l\,.$$

II. eset. A bázist elhagyó vektor a 0. – lépcsőfokhoz tartozik.



Ha a bevonandó vektor is a 0. lépcsőfokhoz tartozik, akkor a két vektor felcserélése a bázisfelépítést nem módosítja, csak a Z_0 blokk változik.

Ha a bevonandó vektor a j-edik lépcsőfokhoz tartozik, végezzük a báziscserét úgy, hogy a bevonandó vektorral X_j elemeinek számát növeljük, Z_0-ból pedig a bázisból kilépő vektort hagyjuk el, a bázisban a két vektor között elhelyezkedő vektorok mindegyikét tegyük a bázis tőle eggyel jobbra lévő oszlopába. A bázisfelépítés nem változik, a blokkok I.1/b-hez hasonlóan módosulnak.

Megjegyzés: (1)

(5) együtthatómátrixának tetszőleges $\hat{B}$ bázisában az l-edik lépcsőfokhoz tartozó vektorok olyanok, hogy a nem-nulla komponenseikből alkotott vektorok (8) egy bázisát tartalmazzák minden $l = 1, 2, \ldots, k$ esetén.

Ha ugyanis a $\hat{B}$ bázis olyan, hogy l_0-adik lépcsőfokhoz tartozó vektorai nem ilyenek, akkor jelöljék ezeket:

$$\begin{pmatrix} \underline{0}_1 \\ \underline{a}_{i_1} \\ \underline{0}_2 \end{pmatrix}, \begin{pmatrix} \underline{0}_1 \\ \underline{a}_{i_2} \\ \underline{0}_2 \end{pmatrix}, \ldots, \begin{pmatrix} \underline{0}_1 \\ \underline{a}_{i_r} \\ \underline{0}_2 \end{pmatrix},$$

ahol: $\underline{0}_1 \in R_{(l_0-1)n_0}$, $\underline{0}_2 \in R_{(k-l_0)n_0+n}$, $\underline{a}_{i_j} \in R_{n_0}$, $\underline{a}_{ij}$ (8) bizonyos együttható vektorai.

Feltevésünk szerint van olyan $\underline{a}$ vektor (8) együtthatóvektorai között, mely nem állítható elő az $\underline{a}_{i_1}, \ldots, \underline{a}_{i_r}$ vektorok lineáris kombinációjaként. De akkor a

$$\begin{pmatrix} \underline{0}_1 \\ \underline{a} \\ \underline{b} \end{pmatrix}$$

vektor nem állítható elő a $\hat{B}$-beli vektorok lineáris kombinációjaként, bármi legyen is a $\underline{b} \in R_{(k-l_0)n_0+n}$ vektor. Ez ellentmondás.

III.

Az (5) feladat megoldására így a következő algoritmust alkalmazhatjuk:

(1) Határozzuk meg a (8) egy B_0 megengedett bázisát, legyen a hozzátartozó bázismegoldás $\underline{x}_0$. Ha (8)-nak nincs megengedett bázisa, fejezzük be az eljárást.

(2) Definiáljuk a B_i bázisokhoz tartozó vektorok indexeinek halmazait $i = 1, 2, \ldots, k$. Minden ilyen indexhalmaz tartalmazza most a B_0-beli vektorok indexeit, és

$$B_i^{-1} = B_0^{-1}, \quad i = 1, 2, \ldots, k.$$

(3) Legyen L_0 a $(T', -T', 0)$ mátrix B_0-beli vektorokkal azonos indexű vektoraiból alkotott mátrix.

$\underline{t} = L_0 \underline{x}_0 - \underline{c}$ felhasználásával határozzuk meg I^*-ot. Definiáljuk a (ρ_1', ρ_2') vektort:

ha I^* tartalmaz negytív egységvektort, akkor $\underline{\rho}_1 = \underline{0}, \underline{\rho}_2$ néhány komponense -1 és első fázisban vagyunk. Egyébként

$$\underline{\rho}_2 = \underline{0}\,, \quad \underline{\rho}_1' = (p_1 \underline{f}_{1,B_0}', \ldots, p_k \underline{f}_{k,B_0}')$$

ahol f_{i,B_0}' a $(\underline{f}_i', -\underline{f}_i', \underline{0}')$ vektor B_0-beli vektorokkal azonos indexű komponenseit tartalmazza; $i = 1, 2, \ldots, k$.

(4) $(LX - Z)^{-1} = -I^*$

(5) $\underline{\pi}_2 = \underline{\rho}_2\,, \quad \underline{\pi}_{1i} = p_i \underline{\rho}_2 L_0 B_0^{-1}$.

(6) $(\underline{\pi}_1, \underline{\pi}_2)$-vel (16), (17), (18), (19) szerint végezzünk optimalitás vizsgálatot.

Első fázis esetén:
(16)-ban $\underline{f}_i = \underline{0}\,, \quad i = 1, 2, \ldots, k$.
(18)-ban $\underline{b} = \underline{0}$ szerepel, és ha az optimalitás kritérium teljesül: módositsuk $(\underline{\rho}_1', \underline{\rho}_2')$-t és folytassuk az eljárást 10-nél.

Második fázis esetén, ha az optimalitási kritérium teljesül: határozzuk meg az optimális megoldást (26) alapján és fejezzük be az eljárást.

(7) Határozzuk meg a bázisba vonandó vektor bázisbeli előállítását (22), (23), (24) illetve (25) felhasználásával. Ha $(\underline{d}_1', \underline{d}_2') \leqslant \underline{0}'$ fejezzük be az eljárást.

(8) (26)-tal határozzuk meg az $(\underline{x}_{B1}', \underline{x}_{B2}')$ megoldásvektort.

(9) Végezzük el a báziscserét és $(LX - Z)^{-1}$ ill. B_i^{-1}, a megfelelő indexhalmazok és $(\underline{\rho}_1', \underline{\rho}_2')$ módosítását c) alapján.

(10) Határozzuk meg $(\underline{\pi}_1', \underline{\pi}_2')$-t (15) felhasználásával.
Folytassuk az eljárást (6)-nál.

Irodalom

[1] Dantzig, G.B. and Mandasky, A: On the solution of Two-stage Linear Programs under Uncertainty, The RAND Corporation paper, P-2039, 28 July, 1960.

[2] Bakes, M.D.: Solution of Special Linear Programming Problems with Additional Constraints, Naval Research Logistics Ouartely 1970. Vol. 17. Nr.4. 411-429 p.

[3] Dantzig–Van Slyke: Generalized Upper Bounding Techniques, Journal of Computer and System Sciences, 1(1967).

[4] Hadley: Linear Programming, Addison-Wesley Publishing Company

[5] Kaul, R.N.: An Extension of Generalized Upper Bounded Techniques for Linear Programming, Operation Research Center University of California 65-27.

[6] Hartmann, K.J. and Lasdon, L.S.: A Generalized Upper Bounding Method for Doubly Coupled Linear Programs.

Summary

ON AN ALGORITHM FOR THE SOLUTION OF THE TWO STAGE STOCHASTIC PROGRAMMING PROBLEM

An algorithm is presented for the solution of the two stage Stochastic Programming problem, when the random variable has a finite number of possible values. In this case the problem reduces to a linear programming problem, and the dual of this is solved using the simplex algorithm. The problem has a very special form, and this gives the idea of specializing the simplex algorithm. This anables us to solve the problem even in case of a high number of values of the random vector.

Резюме

ОБ ОДНОМ МЕТОДЕ РЕШЕНИЯ ПРОБЛЕМЫ ДВУХСТЕПЕННОГО СТОХАСТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

В статье излагается метод решения проблемы двухстепенчатого стохастического программирования Данцига-Маданского при условии, если вероятностный конечный ряд векторного изменения может принимать возможные величины. Процесс для эту задачу специализированный вариант симплексного метода, базовое разложение применяемое на двойное проблемы линейного программирования возникшей в этом случае.

Beérkezett: 1973. szeptember 25.

MTA Számítástechnikai és Automatizálási Kutató Intézete, **Közlemények** 11. (1973)

RITKA MÁTRIXOK INVERTÁLÁSA

dr. Gergely József

BEVEZETÉS

A műszaki gyakorlatban felmerül olyan mátrixok invertálásának igénye, amelyek kevés nem 0 elemet tartalmaznak. Ezekben az esetekben nem célszerű az általános invertálási eljárásokat használni, mert sok felesleges munkát végzünk azzal, hogy a sok 0 elemmel is úgy számolunk, mintha azok 0-tól különbözőek lennének. Ritka (vagy sparse) mátrixok invertálására, valamint ilyen mátrixú lineáris egyenletrendszerek megoldására nagyon sok módszert kidolgoztak. Az [1] és [2] ezzel a témával foglalkozó konferencia anyagait tárgyalja és mintegy 200 irodalmi hivatkozást tartalmaz. Jelen dolgozatban egy ismert képlet ritka mátrixok esetében való célszerű használatát ismertetjük.

MÓDSZER

a) Legyen ismert az $n \times n$-es A mátrix inverze $A^{-1} = D$. Módosítsuk A-t egy $\underline{u}\,\underline{v}^T$ diáddal ($\underline{u}$ oszlop, $\underline{v}^T$ sorvektor). [1]-ben, [3]-ban és még sok irodalomban található a jól ismert következő összefüggés:

$$(1) \qquad (A + \underline{u}\,\underline{v}^T)^{-1} = D - \frac{1}{1 + \underline{v}^T D \underline{u}} D \underline{u}\,\underline{v}^T D,$$

amit ritka mátrixok esetében célszerűen használhatunk.

Adjunk az A mátrix i-edik sorának j-edik eleméhez b-t. Legyen az $\underline{u}$ vektor i-edik komponense 1, a többi 0, a $\underline{v}^T$ vektor j-edik komponense b, a többi 0, vagyis $\underline{u} = \underline{e}_i, \underline{v}^T = b\underline{e}_j^T$ (ahol $\underline{e}_i$ az i-edik egységvektor). Ebben az esetben (1) a következő speciális alakot ölti:

$$(2) \qquad (A + b\underline{e}_i \underline{e}_j)^{-1} = D - \frac{b}{1 + bd_{ji}} \underline{d}_i \underline{d}_j^T$$

ahol d_{ji} a D mátrix j-edik sorának i-edik eleme, $\underline{d}_i$ az i-edik oszlop $\underline{d}_j^T$ pedig a j-edik sorvektora.

A (2) képlet szukcesszive alkalmazható. Ha a A sparse mátrix fődiagonálisában nem 0 elemek vannak, vehetjük az A_1 kiinduló mátrixnak az A fődiagonálisából álló mátrixot. Ennek inverze (D_1) is diagonális (az eredeti mátrix diagonális elemeinek reciprokaiból áll). Vegyünk ezután egy nem 0 fődiagonálison kívüli elemet és alkalmazzuk (1)-et. (1) nagyon egyszerűen hajtható végre, ugyanis minthogy D_1 diagonálmátrix, $d_{ji} = 0, \underline{d}_i$ és $\underline{d}_j^T$ egyetlen nem 0 elemet tartalmazó vektorok. Ezután vegyünk egy újabb fődiagonálison kívüli nem 0 elemet stb. Az eljárást annyiszor kell ismételni, ahány nem 0 elem van. (1) számolása továbbra is nagyon egyszerű lesz, de minden egyes lépés újabb nem 0 elemeket hozhat be a kialakuló inverzbe, az inverz telítetté válhat.

A számolás közben szingularitás lép fel, ha (1)-ben $1 + bd_{ji} = 0$ lesz. Ekkor viszont módunkban áll más nem 0 elemet választani, esetleg ideiglenesen egy nem 0 elemet felvenni, amit később visszaalakíthatunk 0-vá.

Ha az A mátrix fődiagonálisában 0 is van, akkor ezek helyett A_1-be nem 0 elemeket veszünk fel, majd egy későbbi lépésben visszaalakíthatjuk a 0-t.

b) *Több nem 0 elem egyidejű választása*

Legyenek az A i-edik sorában álló $j_1, j_2, \ldots, j_k$ oszlopindexű elemek 0-tól különbözőek, legyenek $b_1, b_2, \ldots, b_k$, a többi 0. Ekkor $\underline{u} = \underline{e}_i$ míg

$$\underline{v}^T(0 \ldots 0\underset{j_1}{b_1}0 \ldots 0\underset{j_2}{b_2}0 \ldots \underset{j_k}{b_k}0 \ldots 0) \text{ alakú lesz.}$$

(1) jobboldalán álló tört ebben az esetben

$$\frac{1}{1 + \underline{v}^T D \underline{u}} = \frac{1}{1 + \sum_{s=1}^{k} b_s d_{j_s i}}$$

$D \underline{u}\, \underline{v}^T D$ pedig olyan diád lesz, melynek baloldali oszlopvektora $\underline{d}_i$, a jobboldali sorvektora pedig

$$\left(\sum_{s=1}^{k} b_s d_{j_s 1}, \sum_{s=1}^{k} b_s d_{j_s 2}, \ldots, \sum_{s=1}^{k} b_s d_{j_s n} \right),$$

vagyis a D mátrix $j_1, j_2, \ldots, j_k$-adik sorának $b_1, b_2, \ldots, b_k$ együtthatókkal vett lineáris kombinációja. E szerint az A mátrix sorainak fődiagonálison kívüli nem 0 elemei egyidejűleg számításba vehetők.

Több nem 0 egyidejű választása egyszerűsíti a számolást, mert soronkint csak egy osztást és egy szingularitás vizsgálatot kell végezni, míg különben annyiszor, ahány nem 0 elem van.

Az ismertetett módszer számológépeken könnyen programozható, amit a CDC 3300-ra el is végeztünk. A számolásba bevonandó nem 0 elemek sorrendje az inverzképzésnél tetszőlegesen választható, amivel a szingularitási problémák elkerülhetők.

Egy nem 0 elem vagy egy sorban lévő nem 0 elemek bevonása a számolásba a kialakuló inverzet egy diád szorzattal módosítja. A diád komponenseit(legalább is a számolás elején) sok 0 elemet tartalmaznak, ami a diádképzés során felhasználható.

ALGORITMUS A b) ESETRE

Az egy sorban álló nem 0 elemeknek a számolásba való bevonásánál célszerű a sorok növekvő sorrendjét követni. Vegyük figyelembe, hogy a már meglévő inverzet, D_k-t módosító diád baloldali oszlopvektora a D_k egy oszlopa. Ha az első sor nem 0 elemeivel kezdünk, akkor az első lépésben $i = 1$ és D_1 első oszlopának ($\underline{d}_1$-nek) csak az első eleme nem 0, így a diád egy sora (első) lesz 0-tól különböző és csak azoknál az oszlopoknál, amelyek indexe megegyezik a bevonandó nem 0 elemek oszlop indexeivel. A második lépésben $i = 2$ és $\underline{d}_2$ első és második eleme különbözhet 0-tól stb. A diád jobboldali sorvektora is erősen hézagos lesz (legalább is a számolás kezdetén) mert a D_k sorainak lineáris kombinációiból képződik, amik viszont csak 1 nem 0 elemet tartalmaznak (a diádikus elemet) mindaddig amíg az eredeti mátrix azonos indexű sorában szereplő nem 0 elemek a számolásba nem kerültek.

MEGJEGYZÉS

A módszer és annak különösen az a) változata célszerűen alkalmazható, ha egy ismert inverzű mátrix egy, vagy kevés számú elemét kell változtatni.

Irodalom

[1] Ralph A. Willoughby: Procedings of the Symposium on Sparse Matrices and Their Applications, Held at the IBM Watson Research Center September 9-10 1968.

[2] J. K. Reid: Large sparse sets of linear equations Proceedings of the Oxford conference; April 1970. Academic press, London 1971.

[3] Faddeev, D. K., and Faddeeva, V. N.: "Vücsiszlityelnüe metodü linejnoj algebrü, Leningrad 1963.

Beérkezett: 1973. október 1.

Summary

INVERSION OF SPARSE MATRICES

It is suggested to apply the relation (1) to the inversion of sparse matrices. The paper analyses case a) when the formula is used by inserting of a non-zero element and case b) when we use the relation (1) by inserting of more non-zero elements from the same row.

Р е з ю м е

ОБРАЩЕНИЕ РЕДКИХ МАТРИЦ

Для обращения редких матриц целесообразно использовать выражение /I/. Статья разбирает случаи употребления этого выражения:

а/ при вводе одного ненулевого элемента и

б/ при вводе одновременно в одну строку нескольких ненулевых элементов.

MTA Számítástechnikai és Automatizálási Kutató Intézete, **Közlemények** 11. (1973)

A FOLYTONOS ÉS DISZKRÉT IDEJŰ STACIONÁRIUS GAUSS-MARKOV FOLYAMAT KAPCSOLATÁRÓL

Krámli András

Arató Mátyás 1969-ben – doktori disszertációjának egy eredményével kapcsolatban – megkérdezte a szerzőtől (és több más matematikustól), hogy a diszkrét és folytonos idejű stacionárius Gauss-Markov folyamat közötti kapcsolatot leíró, a spektrálsűrűség függvényekre vonatkozó analitikus összefüggés igazolható-e egyszerűen, a valószínűségszámítás nélkül is. A keresett bizonyítás a probléma pontos megfogalmazása után néhány sorban leírható:

A folytonos idejű elemi Gauss folyamatok szimulációjával és statisztikájával kapcsolatban felmerült az a kérdés, hogy a $\xi(t)$ autoregressziós folyamat diszkrét idejű megfigyeléseinek $\xi(n)$ sorozata milyen törvényszerűségeknek tesz eleget.

A megoldást az az egyszerű – Doob [2] ismert tételén alapuló – észrevétel szolgáltatta (l. Arató [1]), hogy a folytonos idejű elsőrendű autoregressziós vektorfolyamat "diszkrétizáltja" is elsőrendű autoregressziós folyamat – mert ez folytonos és diszkrét időben egyaránt szükséges és elégséges feltétele a stacionárius Gauss-Markov tulajdonságnak.

A magasabbrendű autóregressziós folyamat diszkrétizáltját ennek az észrevételnek az alapján könnyen kifejezhetjük, mert egy k dimenziós n-edrendű autoregressziós folyamat $k \cdot n = m$ dimenziós elsőrendűként fogható fel. A megfelelő diszkrét idejű folyamat általában nem tiszta autóregressziós hanem autóregressziós-mozgóátlag típusú lesz. Az autóregressziós és mozgóátlag típusú stacionárius Gauss-folyamatok a spektrál elmélet segítségével egyértelműen jellemezhetők: a spektrálsűrűségfüggvényük racionális függvénye λ-nak $(-\infty < \lambda < \infty)$ a folytonos idejű esetben ill. φ-nek $(0 \leqslant \varphi \leqslant 2\pi)$ a diszkrét idejű esetben.
A továbbiakban az egyszerűség kedvéért csak az egydimenziós esetet vizsgáljuk.

Ismeretes, hogy a spektrálelmélet alapján a $\xi(t)$ stacionárius Gauss folyamat a második momentumok szempontjából azonosítható az $e^{it\lambda}\, g(h)$ függvénnyel:

$$R(\xi(t), \xi(s)) = \int_{-\infty}^{\infty} e^{i(t-s)\lambda} |g(\lambda)|^2 d\lambda$$

– ahol $|g(\lambda)|^2 = \tilde{g}(\lambda)$ a spektrálsűrűségfüggvény. Hasonló a $\xi(n)$ folyamat azonosítható $e^{in\varphi} f(\varphi)$-vel, ahol $|f(\varphi)|^2 = \tilde{f}(\varphi)$ a spektrálsűrűségfüggvény.

Minthogy a $\xi(t)$ és a $\xi(n)$ folyamat t egész értékeire ugyanaz és $e^{in\varphi}$ periódikus 2π periódussal, a $[0, 2\pi]$ szakaszon érvényes a következő összefüggés

$$\int_0^{2\pi} e^{i(n-m)\varphi} |f(\varphi)|^2 d\varphi = \sum_{k=-\infty}^{\infty} \int_0^{2\pi} e^{i(n-m)\lambda} |g(\lambda + 2k\pi)|^2 d\lambda$$

azaz – felhasználva a Fourier-sor előállítás egyértelműségét:

$$\widetilde{f}(\varphi) = |f(\varphi)|^2 = \sum_{k=-\infty}^{\infty} |g(\varphi + 2k\pi)|^2 = \sum_{k=-\infty}^{\infty} \widetilde{g}(\varphi + 2k\pi)$$

Az $\widetilde{f}(\varphi)$ ill. $\widetilde{g}(\lambda)$ függvények konkrét alakját behelyettesítve és felhasználva az ismert – Cauchytól származó – komplex függvénytani tételt (l. [3].) azonnal leolvasható, hogy a $\sum_k \widetilde{g}(z + 2k\pi)$ (z most már komplex változó) összeg az $\widetilde{f}(z)$ függvény parciális törtekre való felbontása:

$$\frac{1}{(1 - e^{a+iz})(1 - e^{a-iz})} = \sum_{k=-\infty}^{\infty} \frac{1}{(a + i(z + 2k\pi))(a - i(z + 2k\pi))}$$

A többdimenziós esetben meromorf mátrixfüggvényekre kell megkonstruálni a parciális törtekre való felbontás analogonját, de jelen dolgozatnak csak az alapgondolat ismertetése volt a célja.

Irodalom

[1] Arató M.: Az elemi Gauss folyamatok statisztikai problémái doktori disszertáció (1969)

[2] Doob J. L.: The elementary Gaussian processes, Ann. of Math. Stat. 15, 229-281.

[3] Szőkefalvi-Nagy Béla: Komplex függvénytan, egyetemi jegyzet.

1973. október 15.

Summary

In this note there is given a simple analytic proof of a theorem due to M. Arató, which describes the connection between the continuous time parameter elementary Gaussian process and its discrete time parameter sample-process.

Р е з ю м е

В настоящем заметке дается простое аналитическое доказательство теоремы М. Арато, описывающей связь между элементарным гауссовским процессом с непрерывным временем и процессом дискретных наблюдений того же процесса.

MTA Számítástechnikai és Automatizálási Kutató Intézete, **Közlemények** 11. (1973)

EGY GAUSS-FOLYAMAT MARKOVITÁSÁRÓL ÉS REGULARITÁSÁRÓL

Pergel József

Hannan említi [1] könyvében, hogy az

$$X_n = -\rho^{-1} \sum_{i=0}^{\infty} \rho^{-1} \epsilon_{i+n} \tag{1}$$

folyamat, ahol $|\rho| > 1$, ϵ_i $(i = \ldots -1, 0, 1, \ldots)$ független, normális valószínűségi változók sorozata 0 várható értékkel és 1 szórással, stacionárius megoldása a $\xi_n = \rho\xi_{n-1} + \epsilon_n$ autóregressziós egyenletnek, amely azonban függ az ϵ_i változókban kifejezett ”jövőtől”, azaz ξ_n függ $\epsilon_{n+1}, \ldots, \epsilon_{n+i}, \ldots$ változóktól, ellentétben az ilyen esetekben támasztott szemléletes követelményekkel.

Arató Mátyás vetette fel a kérdést, vajon a fenti folyamat Markov tulajdonságú és reguláris-e. Jelen megjegyzés célja annak bizonyítása, hogy a válasz mindkét kérdésre igenlő. Valójában mindkét állítás bizonyítása igen egyszerű. Ismeretes [2], hogy egy η_n stacionárius Gauss-folyamat akkor és csak akkor Markov tulajdonságú, ha korrelációs függvénye q^n alakú, ahol $|q| < 1$. A fenti X_n stacionárius Gauss folyamat korrelációs függvényét kiszámítva kapjuk, hogy $EX_{t+n} X_t = \rho^{-n}$, tehát X_n Markov tulajdonságú. Más meggondolás is erre az eredményre vezethet. Az $y_n = X_{-n}$ folyamat (X_n időbeni ”megfordítottja”) (1) alapján eleget tesz az

$$y_n = \rho^{-1} y_{n-1} + \rho^{-1} \epsilon_n^1$$

(ahol $\epsilon_n^1 = -\epsilon_{-n+1}$) autóregressziós egyenletnek, ahol ϵ_n^1 független $y_{n-1}, y_{n-2}, \ldots$-tól, tehát y_n Markov tulajdonságú [2]. Másrészt az egydimenziós Gauss folyamat akkor és csak akkor Markov tulajdonságú, ha a ”megfordítottja” is az, ez azonnal következik a korrelációs függvény szimmetriájából.

Ebből az is látható, hogy létezik normális eloszlású valószínűségi változók egy δ_i sorozata $(i = \ldots -1, 0, 1, \ldots)$ 0 várható értékkel és ρ^{-1} szórással, hogy X_n-re fennáll az

$$X_n = \rho^{-1} X_{n-1} + \delta_n$$

autóregressziós egyenlet. δ_n kifejezhető az ϵ_n változók segítségével

$$\delta_n = X_n - \rho^{-1} X_{n-1} = \rho^{-2} \epsilon_{n-1} - \frac{\rho^2 - 1}{\rho^3} \sum_{i=0}^{\infty} \rho^{-i} \epsilon_{i+n}$$

és a δ_n változók függetlensége közvetlen számolással is verifikálható. Hasonló módon fejezhető ki ϵ_n is δ_i változók segítségével. Érvényes tehát az

$$X_n = \sum_{i=0}^{\infty} \rho^{-1} \delta_{n-i}$$

előállítás, bizonyítva X_n regularitását. Ilymódon kimondhatjuk a következő állítást.

Tétel. *A*

$$\xi_n = \rho \xi_{n-1} + \epsilon_n$$

autóregressziós egyenletnek $|\rho| < 1$ *és* $|\rho| > 1$ *esetén is létezik reguláris stacionárius, Gauss megoldása mely Markov tulajdonságú.* $|\rho| < 1$ *esetén a megoldás független a "jövőtől", még* $|\rho| > 1$ *esetén a "múlttól".*

Hasonló meggondolások érvényesek több dimenziós és folytonos idejű stacionárius Gauss-Markov folyamatok esetén is. Többdimenziós esetben ismeretesek példák olyan folyamatokra, amelyek regulárisak, de "megfordítottjuk" szinguláris (lásd pl. [3]). Eredményeinkből látszik, hogy Gauss-Markov esetben ez a helyzet nem állhat elő.

Természetesen többdimenziós esetben a korrelációs függvény szimmetriáján alapuló lépés nem használható, hiszen ebben az esetben nem szimmetrikus a korrelációs függvény. Azonban a közvetlen kiszámolás itt is célravezető. Legyen tehát Q egy k dimenziós mátrix, melynek összes sajátértéke 1-nél nagyobb abszolút értékű, $\underline{\epsilon}_i$ $(i = \ldots 1, 0, 1, \ldots)$ pedig k-dimenziós azonos, normális eloszlású, független valószínűségi változók egy sorozata. Akkor az

(2) $$\underline{\eta}_n = Q\underline{\eta}_{n-1} + \underline{\epsilon}_n$$

egyenletnek megoldása lesz a

(3) $$\underline{Z}_n = -Q^{-1} \sum_{i=0}^{\infty} Q^{-i} \underline{\epsilon}_{i+n}$$

k dimenziós folyamat. (A $\underline{Z}_n$ szorzatot definiáló sor konvergens négyzetes középben és 1 valószínűséggel, mert Q^{-1} sajátértékei Q sajátértékeinek reciprokai, tehát abszolút értékben 1-nél kisebbek). Közvetlen számolással adódik, hogy a

$$\underline{\nu}_n = \underline{Z}_n - \underline{Q}^{-1}\underline{Z}_{n-1} = Q^{-2} \underline{\epsilon}_{n-1} - Q^{-3}(Q^2 - I) \sum_{i=0}^{\infty} Q^{-i} \underline{\epsilon}_{n+i}$$

(ahol I az egység mátrix) sorozat korrelálatlan, tehát, normális lévén, független is. $\underline{\nu}_n$ kovariancia mátrixa

(4) $$E(\underline{\nu}_n \cdot \underline{\nu}_n^*) = Q^{-2} B_{\underline{\epsilon}}\, Q^{-2*} + H$$

alakú, ahol H eleget tesz a

(5) $$Q^{-1} H Q^{-1*} + B_{\underline{\epsilon}} = H$$

egyenletnek és $B_{\underline{\epsilon}}$ pedig $\underline{\epsilon}_n$ kovariancia-mátrixa.

Irodalom

[1] E.J. Hannan: Multiple Time Series, John Wiley and Sons New York (1971.)

[2] Arató M.: Elemi Gauss folyamatok statisztikai problémái.
Doktori disszertáció (1969).

[3] Ю.А. Розанов: Спектральная теория многомерных случайных процессов с дискретным временем, Усп. мат. наук том XIII. выпуск 2I80, (I958).

1973. október 26.

Summary

On the Markovity and regularity of a Gaussian process

The regularity and Markovity of the processes (1) and (3) are proved where $|\rho| > 1$, $|\lambda_i| > 1$ (λ_i are the eigenvalues of Q) Q is a $k \times k$ matrix, $\epsilon_n, \underline{\epsilon}_n$ are 1 and k-dimensional i.i.d. Gaussian random variables. Process (3) is the solution of difference equation (2).

Р е з ю м е

О РЕГУЛЯРНОСТИ И МАРКОВОСКИ НЕКОТОРЫХ ГАУССОВСКИХ ПРОЦЕССОВ

Доказывается релугярность и марковость процессов (I) и (3), где $|\rho| > 1$, $|\lambda_i| > 1$ (λ_i - собственные значения Q), а Q - $k \times k$ матрица, ϵ_n - действительные, $\epsilon_n - k$ - мерные случайные величины с одинаковыми, нормальными распределениями. Эти процессы удовлетворяют стохастическое разностное уравнение (2) в одно и многомерном случае.